

10-029020

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 678.508 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400

Perfect score: 2143

Sequence: 1 MDVKKPKPYRSLTRRRDAER.....EITEDTASSWPVPTDVS LYP 400

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1868@runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2143	100.0	8354	6	Abs52100 Human TEN
2	2138	99.8	3111	2	Aav19251 Human gam
3	2138	99.8	3111	2	Aax87705 Gamma-her
4	2138	99.8	3111	3	Aad00791 Human Her
5	2138	99.8	3111	4	Aas18526 DNA encod
6	2098	97.9	8438	6	Abn85378 Human NOV
7	2082	97.2	8645	6	Abs78652 Human cdn
8	1285	60.0	2387	2	Aav19252 Human gam

9	987	46.1	1534	9	ADE07179	Ade07179 Novel cod
10	987	46.1	8645	6	ABQ82344	Abq82344 Human NOV
11	987	46.1	8675	6	ABQ82343	Abq82343 Human NOV
12	948	44.2	9695	7	ACC72052	Acc72052 BCU0205B
13	839	39.2	9729	5	AAS14089	Aas14089 Human FCT
14	839	39.2	9729	9	ADB32028	Adb32028 Human FCT
15	839	39.2	9826	5	AAS14085	Aas14085 Human FCT
16	839	39.2	9826	9	ADB32023	Adb32023 Human FCT
17	742.5	34.6	13202	4	AAK51828	Aak51828 Human pol
18	734.5	34.3	12879	6	ABK92230	Abk92230 Prostate
19	653.5	30.5	8473	6	ABQ82345	Abq82345 Human NOV
20	653.5	30.5	8487	6	ABQ82346	Abq82346 Human NOV
21	539	25.2	3052	6	ABQ75971	Abq75971 Human pai
22	534	24.9	1755	4	AAF23418	Aaf23418 Human SEC
23	533.5	24.9	978	9	ADE08934	Ade08934 Novel DNA
24	508	23.7	1727	7	ADA53142	Ada53142 Human cod
25	508	23.7	9058	7	ACC72051	Acc72051 BCU0205A
26	434	20.3	1429	8	ACD40264	Acd40264 Human bre
27	434	20.3	1430	5	AAS14084	Aas14084 Human FCT
28	434	20.3	1430	9	ADB32021	Adb32021 Human FCT
29	434	20.3	1431	4	AAF27861	Aaf27861 Human NOV
30	434	20.3	1743	5	AAS01213	Aas01213 DNA encod
31	404	18.9	2676	5	AAS67421	Aas67421 DNA encod
32	344.5	16.1	452	6	ABS72992	Abs72992 Human gen
33	279	13.0	328	2	AAV87144	Aav87144 EST clone
34	279	13.0	1692	7	ACD05836	Acd05836 Novel hum
35	277	12.9	643	5	AAS67419	Aas67419 DNA encod
36	167.5	7.8	784	5	AAS69256	Aas69256 DNA encod
37	142.5	6.6	24533	4	AAS27689	Aas27689 DNA encod
38	142.5	6.6	24533	7	ABT17017	Abt17017 Human sec
39	142.5	6.6	24533	7	ABZ68157	Abz68157 Human sec
40	142.5	6.6	24533	7	ABZ74637	Abz74637 Secreted
41	142.5	6.6	24533	9	ADB94492	Adb94492 Novel hum
42	141	6.6	9222	7	AAL61171	Aal61171 Actinosyn
43	141	6.6	82746	7	AAL61224	Aal61224 Actinosyn
44	140	6.5	6034	9	ADD29677	Add29677 Human tum
45	140	6.5	6131	6	ABZ11281	Abz11281 Human pol

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
XX
AC ABS52100;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human TEN-M4-like gene.
XX

KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophilin;
KW single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers
variation replace(117,G)
/*tag= a


```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
PN 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
XX 20-DEC-2000; 2000US-0257314P.
PR 02-MAY-2001; 2001US-0288153P.
PR 29-MAY-2001; 2001US-0294075P.
PR 24-JUL-2001; 2001US-0307506P.
PR 10-AUG-2001; 2001US-0311590P.
PR 10-AUG-2001; 2001US-0311613P.
PR 29-AUG-2001; 2001US-0315617P.
PR 14-SEP-2001; 2001US-0322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR P-PSDB; ABG70388.
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.
XX
PS Claim 8; Page 50-52; 318pp; English.
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders. The
CC present nucleic acid sequence represents a NOVX gene. This sequence
CC encodes a NOVX protein of the invention
XX
SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.42e-127 Length: 8354
Score: 2143.00 Matches: 400
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x ABS52100 (1-8354)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db ATGGACGTGAAGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCGGACGCGGAGCGC 94
```

```
QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db CGCTACACCAAGCTCGTCCCGCGGACAGCGAGGAGGGCAAGCCCCGAGAAATCGTACAGC 154
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db TCCAGCGAGACCTTGAAGGCGCTACGACCAAGGACGCCCGCTAGCCTATGGCAGCGCGTC 214
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db AAGGACATTGTCCCGCAGGAGCGCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG 274
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db CGGGAGCTGGGGCTGGAAGAGTAACGCCCCCTCACGGGACCTGTACCGACAGACATT 334
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db GGCCTGCCCAATGCGGCTACTCCATGGGGCTGGCTCTGTGTCGACATGGAGGCTGAC 394
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db ACGTGCTGTCCCTGAGCACCCCGTGGCTGTGTGGGCGGAGCACACGGTCAGGGCGC 454
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db AGTCTCGCTGTCCAGCGCGGCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 514
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCAACGCGCGGCTCCGGACGCCGCG 574
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db CCGCGCTCTGCAACGCCCAACACCCCAACAGCACACCGCGGCTCCATTAATCCCTG 634
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCGCGGCGGCGGCGGCGGCGGCGG 694
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db TCCGGAGAGCCCCCTGCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 754
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db CTCACAGCAACATCCCCCTGGAGACCAAGAACCTAGGCAAGCAGCATTCCTAGGGACA 814
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db TTGCAGGACAACCTCATTTAGATGGACATTTCTCGGCGCTCTCCGCGCATGATGGGCTTAC 874
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db AGTGACGGGCACTTCTCTTCAAGCCTGGAGGCACCTCCCGGCTCTTCTGCACCATCA 934
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db CCAGGGTACCCACTGACGTCCAGCAGTGTACTCTCTCCGCGCGGCGGCGGCGGCGGCGG 994
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db AGCACCTTCGCGCGGCGGCGGCTTTAACTCAAGAGCCCTCAAGAGCTCTCAAGTGAAG 1054
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db TGCGCAGCCCTGAGCGCCATCGTCTCATCTCAGCCACTCTGGTCTATCCTGCTGGCATACTTT 1114
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db GTGGCCATGCACCTGTTTGGCCTAACTGGCACCTGCAGCCCGATGGAGGGGCGAGATGTAT 1174
```

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1175 GAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1234

RESULT 2
AAV19251
ID AAV19251 standard; cDNA; 3111 BP.
XX
AC AAV19251;
XX
DT 17-AUG-1998 (first entry)
XX
DE Human gamma-heregulin cDNA.
XX
KW Gamma-heregulin; gamma-HRG; human; autocrine growth factor;
KW breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
KW cell proliferation; cell differentiation; cell survival;
KW neurological disorder; muscular disorder; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 534..2645
FT /*tag= a

XX WO9802541-A1.
PN 22-JAN-1998.
XX
PF 08-JUL-1997; 97WO-US011841.
XX
PR 12-JUL-1996; 96US-0021640P.
XX
PA (GETH) GENENTECH INC.

XX Schaefer GM, Sliwowski M;
DR WPI; 1998-110589/10.
DR P-PSDB; AAW44817.
XX

PT DNA encoding gamma-heregulin - used to activate ErbB receptor and to
PT enhance proliferation, differentiation or survival of a cell.

XX Claim 21; Fig 1A-C; 81pp; English.

CC This nucleic acid molecule codes for human gamma-heregulin (gamma-HRG)
CC (see AAW44817), a novel member of the heregulin superfamily, that has a
CC unique N-terminal domain not present in previously identified heregulins.
CC Gamma-HRG is secreted by human breast cancer MDA-MB-175 cells. It leads
CC to the formation of a constitutive active receptor complex and stimulates
CC the growth of these cells in an autocrine manner. The nucleic acid was
CC isolated by screening a MDA-MB-175 with a probe corresponding to the EGF-
CC like domain and part of the N-terminal sequences of HRG-beta-3. It can be
CC used for the recombinant production of gamma-HRG, or for in vivo or ex
CC vivo gene therapy. A claimed nucleic acid, which is complementary to the
CC nucleic acid sequence encoding the N-terminal domain of gamma-HRG, is
CC able to reduce production of gamma-HRG by MDA-MB-175 cells. The invention
CC provides claimed methods for activating an ErbB receptor and for
CC enhancing proliferation, differentiation or survival of a cell by
CC contacting the cell (preferably a glial or muscle cell) with gamma-HRG

XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.62e-127 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AAV19251 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGACGTGAAGAGAGAGGAGGCTTACCGCTCGCTGACCCCGCGCGGACGCCGAGCGC 393
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCAAGCTCGTCCGGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCTGAAGGCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGCGTC 513
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTGTCCGCGAGGAGCGGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCTCCCCCACTGCGGCTACTTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGTGTCTGTCCCCGTAGCACCCCGTGTGCTGTGGGGCGGAGCACACGGTCAAGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCCTGTCCAGCCGGGCCAAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACCCACGCGGGCTCCGGACCGCCCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTCGCACGCCCAACCCCCAACCCAGCACCCCGCGGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGGCAACTTCACCGCAGGAGCAACCCCAAGCCCGCCCCCACGGACCACTCGTCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCGCCAGGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAAGAACCTTAGSAAAGCAGCCATTTCCTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCTCCCGCCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACCTTCCTCTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGCCCGCTGCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAGAGCCCTCCAAGTACTGTAACTTGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGCTGGCATACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1414 GTGGCCATGCACCTGTTTGGCCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 3
AAX87705
ID AAX87705 standard; cDNA; 3111 BP.
XX
AC AAX87705;
XX
DT 26-OCT-1999 (first entry)
XX
DE Gamma-hereregulin cDNA.
XX
KW Gamma-hereregulin; HRG-alpha; human; ligand; HER2; HER3; HER4; receptor;
KW lung surfactant; respiratory distress syndrome; emphysema;
KW epithelial growth factor; therapy; ss.

OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 334..2640
FT /*tag= a
XX
PN WO9939729-A2.
XX
PD 12-AUG-1999.
XX
PF 03-FEB-1999; 99WO-US002390.
XX
PR 04-FEB-1998; 98US-00020598.
XX
PA (GETH) GENENTECH INC.
PA (IOWA) UNIV IOWA RES FOUND.
XX
PI Sliwkowski M, Kern JA;
XX
DR WPI; 1999-494213/41.
DR P-PSDB; AAY06639.
XX
PT Heregulin ligands can be used to induce epithelial cell growth, and to
PT promote repair and healing of tissue damage or injury.
XX
PS Disclosure; Page 111-116; 120pp; English.

CC This is the nucleotide sequence of cDNA coding for gamma-hereregulin (gamma
CC -HRG, see AAY06639). The invention provides HRG ligands, including gamma-
CC HRG, that have affinity for and stimulate HER2, HER3 and/or HER4
CC receptors in autophosphorylation. A new method of treating respiratory
CC distress syndrome in humans uses HER2, HER3 and/or HER4 receptor ligands
CC as epithelial growth factors. A novel method of inducing epithelial cell
CC growth and/or proliferation comprises contacting a normal epithelial cell
CC which expresses HER2, HER3 and/or HER4 receptors with an isolated ligand
CC which activates HER2, HER3, HER4 receptors or their combination. Also
CC claimed are methods of increasing lung surfactant protein A, or of
CC treating chronic obstructive pulmonary disease, respiratory distress or
CC emphysema, by administering an effective amount of an isolated HER ligand
CC to a patient

SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.62e-127 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AAX87705 (1-3111)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGAGAGAGGAGCCTTACCGCTCGTGACCCGCGCGCGAGCGCGC 393
Qy 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCAGCTCGTCCGCGGACAGCGAGGGGCAAGCCCGCAGAAATCGTACAGC 453
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC 513
Qy 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATGTGCCGACGAGGCGGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG 573
Qy 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGACAGACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCCACTCGGGTACTCCATGGGGCTGGCTCTGTATGCCGACATGGAGGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGTGCTGTCCCCCTGAGCACCCCGTGGCTGTGTGGGCGCGGACACACGGTCAGGGCGC 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGGGCTGCAGAACCCAGCGCGGCTCCGACGCCGCCG 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCCGCTCTCGCAGCGCCACACCCCAACACAGCACACCGCGGCTCCATTAACTCCCTG 933
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCCGAGGAGCAACCCCGCGGCCCGCCACGACCATCGCTC 993
Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACCGCCAGGAGAACTGGCTG 1053
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGAGCCATTCTTAGGGACA 1113
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCATGATGGGGCTTAC 1173
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGCACCTCCCCGCTCTCTGCACCACATCA 1233
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTACCCACTGACGTCCAGCACAGTGACTCTCTCTCCGCGCCGACCCCTGCCCGC 1293
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTyrLys 340
Db 1294 AGCACCTTCGCCCGCGCGCCCTTAACCTCAAGAAGCCCTCCAGTACTGTAACTGGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1354 TGCGCAGCCCTGAGCGCCCATCGTCACTCTCAGCCACTCTGGTCACTCTGTGGCATACTTT 1413

[illegible]

Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTCGCCGCGCCCTTTAACTCAAGAGCCCTCCAAGTACTGTAAGTGAAG 1353
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1354 TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTGCTGGCATACTTT 1413
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCGACGTCTCCCTATACCCC 1533
RESULT 5
AAS18526
ID AAS18526 standard; cDNA; 3111 BP.
XX
AC AAS18526;
XX
DT 26-FEB-2002 (first entry)
XX
DE DNA encoding human heregulin, gamma-HRG.
XX
KW Human; heregulin; antiasthmatic; antiinflammatory; vulnery; antiulcer;
KW epithelial cell growth; HER2; HER3; HER4; HRG; lung cell; asthma;
KW lung surfactant protein A; respiratory distress; emphysema; pneumonitis;
KW chronic obstructive pulmonary disease; COPD; chronic bronchitis; ulcer;
KW neonatal pulmonary disease; neonatal respiratory distress syndrome;
KW meconium aspiration syndrome; congenital diaphragmatic hernia; lesion;
KW acute lung injury; cystic fibrosis; surgical wound; resection;
KW growth factor; smoke inhalation; gamma-HRG; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 334..2640
FT /*tag= a
FT /product= "Human heregulin, gamma-HRG"

XX US2001023241-A1.
PN 20-SEP-2001.
PD 02-FEB-2001; 2001US-00773517.
XX
PF 04-FEB-1998; 98US-0073866P.
PR 02-FEB-1999; 99US-00243198.
XX
PA (SLIW/) SLIWKOWSKI M X.
PA (KERN/) KERN J A.
XX
PI Sliwkowski MX, Kern JA;
XX
DR WPI; 2001-595807/67.
DR P-PSDB; AAU09891.
XX
PT Inducing epithelial cell growth and/or proliferation, useful in the
PT treatment of respiratory disease, comprises use of heregulin ligand as
XX growth factor.
XX
PS Disclosure; Fig 7; 68pp; English.
XX
CC The invention relates to inducing epithelial cell growth and/or
CC proliferation comprising contacting a normal epithelial cell, which
CC expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4,
CC HER3/HER4, HER3 or HER4) receptors with an isolated ligand such as
CC heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The
CC method is used for inducing epithelial cell (preferably lung cell) growth
CC and/or proliferation, for increasing lung surfactant protein A and for
CC treating respiratory distress or emphysema, for treating chronic

CC obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma,
CC neonatal pulmonary diseases including neonatal respiratory distress
CC syndrome, meconium aspiration syndrome, chronic lung disease of the
CC neonate, congenital diaphragmatic hernia and acute lung injuries
CC including smoke or chemical inhalation, pneumonitis due to aspiration,
CC radiation, near drowning, cystic fibrosis and other epithelial cell
CC trauma diseases including injuries associated with surgical wounds and
CC resections, ulcers, lesions and tissue tears as normal epithelial cell
CC growth factors. For treating infants/neonates with respiratory distress
CC as well as youth and adult with poor lung function due to lung injury or
CC damage. The HRG binds with varying and very high affinity to the HER2,
CC HER3 and/or HER4 receptors. The method stimulates growth and
CC proliferation of the epithelial cells, repairing and re-establishing the
CC cellular barriers of organs and allowing the affected tissue to develop
CC normal physiological functions more quickly. Hence the method improves
CC oxygenation and speeds of the development of a barrier to infection while
CC treating lung cells that are damaged by inhalation of smoke resulting in
CC emphysema. It also facilitates regeneration of epithelial cells. The
CC present sequence represents the coding sequence of human heregulin gamma-
CC HRG
XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.62e-127 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 4 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x AAS18526 (1-3111)
Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGGAGAGGAGGAGCCTTACCGCTCGCTGACCCGCGCGCGAGCGCGC 393
Qy 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACAGCTCGTCCGCGGACAGCGAGGAGGGCAAGCCCCGAGAAATCGTACAGC 453
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCTGAAGGCCTACGACGAGGACGCGCCCGCTAGCTATGGCAGCGCGTC 513
Qy 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTGTCCGCGAGGAGGCCGAGGAATTCGCCGACAGGTGCCAACTTACCCCTG 573
Qy 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCGGACATGGAGGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCCGTGGTCTGTGGGCGCGGAGCACACGGTACGGGCGC 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCCTGCCTGTCCAGCCGGGCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGGCGCCTGCAGAACCCACCGCGGGCTCCGAGCGCCCG 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCCGCTCTCGCAGCGCCCAACACCCCAACAGCACCGCGGGCTCCATTAACTCCCTG 933

Db 364 ACGTGCTGTCCCTTGAGCACCCCGTGGTCTGTGGGGCCGGAGCACCGTTCAGGGCGC 423
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 424 AGCTCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTACCGACACCGAGCATGAA 483
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 484 AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACCAACGCGCGCTCCGGACGCGCGC 543
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 544 CCGCGCTCTCGCACGCCACACCCCAACACGAGCACACCGCGGCTCCATTAACCTCCCTG 603
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 604 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGCCCGCCACGACCACTCGCTC 663
Qy 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 664 TCCGGAGAGCCCCCTGCGCGCGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG 723
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 724 CTCAACAGCAACATCCCCCTGGAGACCAAGAACCTAGGCAAGCAGCCATTCCTAGGGACA 783
Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 784 TTGCAGGACAACCTCATTTGAGATGGACATTCGCGCGCTCCCGCATGATGGGCTTAC 843
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 844 AGTGACGGGCACCTCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCA 903
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 904 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCGCCGACCCCTGCCCGC 963
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 964 AGCACCTTCGCTGGCGCGCTTTAACTCAAGAGCCCTCCAAGTACTGTAACTGGAG 1023
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1024 TGCGAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCACTCTGTGGCATACTT 1083
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
Db 1084 GTGGCATGCACCTGTTTGGCCTAACTGGCACCTGCAGCCGATGGAGGGCAGATGACG 1143
Qy 380 ----TyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
Db 1144 GATTATGAGATCACGGAGGACACAGCCAGCAGTGTGGCCTGTGCCAACCGACGTCCTCCT 1203
Qy 398 uTyrPro 400
Db 1204 ATACCCC 1210
RESULT 7
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
AC ABS78652;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;

KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200272830-A2.
XX
PD 19-SEP-2002.
XX
PF 08-FEB-2002; 2002WO-US003715.
XX
PR 09-FEB-2001; 2001US-0268111P.
PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
PR 09-MAR-2001; 2001US-0274552P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
DR WPI; 2002-723356/78.
DR P-PSDB; ABG97359.
XX
PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX
PS Claim 5; Page 175-178; 181pp; English.
XX
CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGD polynucleotides and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores: 2.02e-123 Length: 8645
Pred. No.: 2082.00 Matches: 393
Score:

Percent Similarity: 98.25% Conservative: 0			
Best Local Similarity: 98.25% Mismatches: 3			
Query Match: 97.15% Indels: 4			
DB: 6 Gaps: 2			
US-10-029-020-14_COPY_1_400 (1-400) x ABS78652 (1-8645)			
Qy	1	MetAspValIysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	117	ATGGACGTGAAGAGAGGAAGCCTTACCGCTCGCTGACCCGCGCGGACGCCGAGCGC	176
Qy	21	ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	177	CGCTACACCAGCTCGTCCGCGGACAGCGAGGAGGCAAGCCCCGAGAAATCGTACAGC	236
Qy	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	237	TCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGCGTC	296
Qy	61	LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu	80
Db	297	AAGGACATTGTGCCGCGAGGAGGCCGGAATTCTGCCGCACAGGTGCTCAACCTTCACCCCTG	356
Qy	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	357	CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT	416
Qy	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	417	GGCCTCCGCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	476
Qy	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140
Db	477	ACGGTGCTGTCCCTTGAGCACCCCGTGGCTGTGTGGGCGCGGACACCGTCCAGGGCGC	536
Qy	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	537	AGCTCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA	596
Qy	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	597	AACACTGAGACT-----CCGGGGCGCTGCAGAACCCAGCGGGCTCCGGACGCGCGCG	650
Qy	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200
Db	651	CGCGCGCTCTCGACGCCCCACACCCCAACCCAGCACCCAGCGGGCTCCATTAACCTCCCTG	710
Qy	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	711	AACCGGGCAACTTCACGCGGAGAGCAACCCAGCCCGCGCCACCGACCACTCGCTC	770
Qy	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	771	TCCGAGAGCCCCCTGCCGGCGCGCCAGGAGCCTGCCACGCCCCAGGAGAACTGGCTG	830
Qy	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr	260
Db	831	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA	890
Qy	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	891	TTGCAGGACAACCTCATTTGATGGACATTCTCGGGCGCTCCCGCCATGATGGGGCTTAC	950
Qy	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	951	AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCACATCA	1010
Qy	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg	320
Db	1011	CCAGGGTACCCCACTGACGTCCAGCACAGTGTACTCTCTCTCCGCCCGACCCCTGCCCGC	1070
Qy	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340

Db	1071	AGCACCTTCGCCCCGGCGCCTTTAACTCTCAAGAGCCCTCCAAGTACTGTAACTGGAAG	1130
Qy	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1131	TGGCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTT	1190
Qy	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1191	GTGGGTAAGCACCTCTTC-----AACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT	1244
Qy	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	1245	GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC	1304
RESULT 8			
AAV19252			
ID	AAV19252	standard; cDNA; 2387 BP.	
XX	AAV19252;		
AC	AAV19252;		
XX	17-AUG-1998	(first entry)	
DT	17-AUG-1998	(first entry)	
XX	Human gamma-heregulin cDNA clone 20.		
DE	Human gamma-heregulin cDNA clone 20.		
XX	Gamma-heregulin; gamma-HRG; human; autocrine growth factor;		
KW	breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;		
KW	cell proliferation; cell differentiation; cell survival;		
KW	neurological disorder; muscular disorder; ss.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FH	CDS	307..1890	
FT		/*tag= a	
FT	misc_feature	1186..1263	
FT		/*tag= b	
FT		/note= "insert DNA"	
XX	WO9802541-A1.		
PN	22-JAN-1998.		
XX	08-JUL-1997;	97WO-US011841.	
PF	12-JUL-1996;	96US-0021640P.	
XX	(GETH) GENENTECH INC.		
PA	Schaefer GM, Sliwowski M;		
XX	WPI; 1998-110589/10.		
XX	P-PSDB; AAW44818.		
DR	DNA encoding gamma-heregulin - used to activate ErbB receptor and to		
DR	enhance proliferation, differentiation or survival of a cell.		
XX	Disclosure; Page 56-57; 81pp; English.		
PS	Clone 20 encodes an N-terminally truncated isoform (see AAW44818) of		
XX	human gamma-heregulin (gamma-HRG) (see also AAW44817), a novel member of		
CC	the heregulin superfamily. It was isolated from a MDA-MB-175 human breast		
CC	cancer cell library after screening with an EGF-like domain. Compared to		
CC	gamma-HRG cDNA (see AAV19251), clone 20 contains a 78 base insert between		
CC	codons 560 and 561, and lacks the 5' end. The truncated gamma-HRG was		
CC	insoluble when expressed as a thioiodoxin fusion protein in Escherichia		
CC	coli cells. Gamma-HRG nucleic acid and polypeptide are claimed, together		
CC	with their uses e.g. for enhancing the proliferation, differentiation or		
CC	survival of glial or muscle cells		
XX	Sequence 2387 BP; 641 A; 640 C; 550 G; 556 T; 0 U; 0 Other;		
SQ	Alignment Scores:		
	Pred. No.:	4.66e-73	Length: 2387

Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.96% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AAV19252 (1-2387)

Qy 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeuSerHis 185
|||
Db 1 CATCCGGCGGCGCTGCAGAACACCGCGGGCTCCGGACGCGCGCGCGCTCTCGCAC 60
|||

Qy 186 AlaHisThrProAsnGlnHisHisAlaSerIleAsnSerIleAsnArgGlyAsnPhe 205
|||
Db 61 GCCCACACCCCAACACGACACACGCGGCTCCATTAACCTCCCTGAACCGGGCACTTC 120
|||

Qy 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
|||
Db 121 ACGCCGAGGAGCAACCCAGCCCGGCCACGGACCACTCTCGGAGAGCCCT 180
|||

Qy 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
|||
Db 181 GCCGGCGGCGCCAGAGCCTGCCACGCCAGGAGAACTGGCTGTCAACAGCAACATC 240
|||

Qy 246 ProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThrLeuGlnAspAsnLeu 265
|||
Db 241 CCCCTGGAGACCAAGAACCTAGGCAAGCAGCCATTCCTAGGGACATTGCAGGACAACCTC 300
|||

Qy 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
|||
Db 301 ATTGAGATGGACATTCTCGGGCCCTCCCGCCATGATGGGCTTACAGTGACGGGCATTC 360
|||

Qy 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
|||
Db 361 CTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCACATCACAGGGTACCCACTG 420
|||

Qy 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325
|||
Db 421 ACGTCCAGACAGTGTACTCTCTCCGCGCCCGACCCCTGCCCCGAGCACCTTCGCCCGG 480
|||

Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaLeuSer 345
|||
Db 481 CCGGCTTTAACTCAAGAGCCCTCCAGTACTGTAACTGGAAAGTGGCAGCCCTGAGC 540
|||

Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPheValAlaMetHisLeu 365
|||
Db 541 GCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTTGTGGCCATGCACCTG 600
|||

Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
|||
Db 601 TTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTATGAGATCACGGAGGAC 660
|||

Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
|||
Db 661 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 705
|||

RESULT 9
ADE07179
ID ADE07179 standard; DNA; 1534 BP.
XX
AC ADE07179;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel coding sequence (useful for identifying genetic disorders) #245.
XX
KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder; gene; ds.
XX
OS Unidentified.
XX
PN WO2003054152-A2.
XX

PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002WO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
DR WPI; 2003-569235/53.
DR P-PSDB; ADE08090.
XX
PT New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
PS Claim 1; SEQ ID NO 245; 1177pp; English.
XX
CC The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present DNA sequence represents a gene of the
CC invention.
XX
SQ Sequence 1534 BP; 462 A; 377 C; 355 G; 340 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.3e-54 Length: 1534
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best Local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 9 Gaps: 12

US-10-029-020-14_COPY_1_400 (1-400) x ADE07179 (1-1534)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
|||
Db 133 ATGGATGTGAAAGAACGCGAGCCCTTACTGCTCCCTGACCAAGACGACGAGAGGAA 192
|||

Qy 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
|||
Db 193 CGGCGCTACACAAATTCTCCGACAGCAATGAGGAGTGCCGGGTACCCACACAGAAGTCC 252
|||

Qy 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
|||
Db 253 TACAGTTCGACGAGACATTGAAAGCTTTTGATCATGATTCTCGCGGCTGCTTTACGGC 312
|||

Qy 58 SerArgValLysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
:::|
Db 313 AACAGAGTGAAGGATTGGTTTCACAGAGAAGCAGACGAGTTTCATAGACAAGACAGAAT 372
|||

Qy 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
|||
Db 373 TTTACCCCTAAGGAGTTAGGAGTTTGTGAACCCAGCAACTCGAAGAGGACTGGCATTGT 432
|||

Qy 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
:::|
Db 433 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCATGTCAGGGTCAGATGCTGATACT 492
|||

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 493 GAAATGAAGCAGTGATGTCCCCAGAGCATGCCATGAGACTTTGGGCAGGGGTCAA 552
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 553 TCAGGCGCAGCTCCTGCCTGTCAAGTCGGTCCAACCTCAGCCCTCACCCCTGACAGATACG 612
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 613 GAGCAGCAAAACAAGTCGAGTGAGAAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 672
QY 175 ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 673 ACCCTGAGCCCTTGCCGCCT-----TCCATAAGCAGCAGCTCTGCACAGCATCAT--- 723
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 724 CCATCCCATCACTTCTCTCAACAGAAACTCCCTGACCATAAGAGGAACCCAGAGTCCGGCC 783
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 784 CCG-----CCGGCTGCTTTGCCCGCGGAGCTGCAAAACCACA 819
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 820 CCCGAGTCCGTCAGCTGCAGGACAGCTGGTCTCTGGCAGTAATGTACCCTGGAAAGC 879
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 880 AGG----- 882
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 883 -----CATTTCTATTCAAAACA 900
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 901 GGAACAGGTACAACGCCACTGTTCAGTACTGCAACCCCGAGATACACAATGGCATCTGCC 960
QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 961 TCTGTTTATTCAACCACTACTCGCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTT 1020
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1021 AAATTCAGAAGTCTTCAAGTACTGTAGCTGGAATGCACTGCACTGTGTGCCGTAGGG 1080
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1081 GTCTCGGTGCTCCTGGCAATACTCCTGTCTTATTTATAGCAATGCATCTCTTTGGCCTC 1140
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1141 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGAAATGGAAGTGAATTCTGAT 1200
QY 386 ThrAlaSerSerTrpProValProThrProThrAspValSerLeuTyrPro 400
Db 1201 ACC-----ATGCCAACAAACACTGTGTCTTACCT 1230

RESULT 10
ABQ82344

ID ABQ82344 standard; cDNA; 8645 BP.

XX AC ABQ82344;

XX AC ABQ82344;

XX AC ABQ82344;

DE 17-DEC-2002 (first entry)

DE Human NOV15b encoding cDNA SEQ ID NO:37.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;

KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;

KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;

KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disease; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

OS
XX
FH Location/Qualifiers
FT 151..8316
FT /*tag= a
FT /product= "NOV15b"

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.

XX 02-JAN-2001; 2001US-0259415P.

XX 04-JAN-2001; 2001US-0259785P.

XX 20-FEB-2001; 2001US-0269814P.

XX 09-MAR-2001; 2001US-0279863P.

XX 29-MAR-2001; 2001US-0279832P.

XX 13-APR-2001; 2001US-0283889P.

XX 18-APR-2001; 2001US-0284447P.

XX 25-APR-2001; 2001US-0286683P.

XX 29-MAY-2001; 2001US-0294080P.

XX 16-AUG-2001; 2001US-0312915P.

XX 17-AUG-2001; 2001US-0313325P.

XX 17-SEP-2001; 2001US-0322699P.

XX 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

XX Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

XX Kekuda R, Patturajan M, Gusev V, Gangolli RA, Guo X, Shenoy S;

XX Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

XX Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.

XX P-PSDB; ABP53587.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-

XX associated disorders, such as cancers, neurological disorders, disorders

XX of vesicular transport, gastrointestinal disorders, and autoimmune

XX diseases.

XX Claim 8; Page 114-117; 444pp; English.

XX The present invention describes novel human proteins designated NOVX,

XX where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,

XX cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,

XX antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,

XX immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
XX protozoacide and antihelminthic activities, and can be used in gene
XX therapy. The NOVX proteins, nucleotides or antibodies can be used in the
XX manufacture of a medicament for treating a syndrome associated with a
XX human disease selected from NOVX-associated disorder, such as cancers
XX (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
XX ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
XX ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
XX disease), disorders of vesicular transport (e.g. cystic fibrosis,
XX diabetes mellitus, Grave's disease, or goitre), gastrointestinal
XX disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
XX autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
XX anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic

PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX WPI; 2002-732706/79.
DR P-PSDB; ABP53586.

XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.

XX
PS Claim 8; Page 110-112; 444pp; English.

XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4

XX
SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	2.43e-53	Length:	8675
Score:	987.00	Matches:	212
Percent Similarity:	62.89%	Conservative:	49
Best Local Similarity:	51.08%	Mismatches:	90
Query Match:	46.06%	Indels:	64
DB:	6	Gaps:	12

US-10-029-020-14_COPY_1_400 (1-400) x ABQ82343 (1-8675)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAGAACGAGCCCTTACTGCTCCCTGACCAAGAGCAGACGAGAAGGAA 210
QY 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACACAAATTCCTCGCAGACAAATGAGGAGTCCGGGTACCCACACAGAAGTCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db

Db 271 TACAGTTCACGCGAGACATTGAAAGCTTTTGTATCATGATTCTCTCGCGGCTGCTTTACGGC 330
QY 58 SerArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGGATTGGTTTCACAGAGAAGCAGACGAGTTTCACTAGACAAGGACAGAAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCCCTAAGGCAGTTAGGAGTTTGTGAACCAAGCAACTCGAAGAGGACTGGCATTTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAATGGGCTCCCTCACAGAGGTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 511 GAAAATGAAGCAGTGATGTCCCCAGAGCATGCCATGAGACTTTGGGGCAGGGGTCAAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGCAGCTCCTGCCTGTCAAGTCGGTCCAACTCAGCCCTCACCTGACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAAACAAGTCCGACAGTGAGAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690
QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGCCGCT-----TCCCATAAAGCAGCACTCTGCACAGCATCAT 741
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCTCAACAGAAACTCCCTGACCAATAGAAAGAACAGAGTCCGGCC 801
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGGCTGCTTTGCCCGCGAGCTGCAAAACCA 837
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCGAGTCCGTCCAGTGCAGCAGTGGTCCCTTGGCAGTAATGTACCACCTGGAAGC 897
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTATTCAAACA 918
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAACGCCACTGTTTCAGTACTGCAACCCAGGATACACAATGGCATCTGGC 978
QY 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTCACCACCTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAAGTCTTT 1038
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAAGAAGTCTTCAAGTACTGTAGCTGGAAATGCACTGCACTGTGTGCGGTAGGG 1098
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCCTGGCAATACTCCTGTCTTATTTATAGCAATGCATCTCTTTGGCCTC 1158
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGAATGGAAAGTGAATTCTGAT 1218
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAAACACTGTGTCACTTACCT 1248

RESULT 12
ACC72052

RECEIVED 12
ACC72052
ID ACC72052 standard: DNA: 9695 BP.

ACC72052;

08-JUL-2003 (first entry)

DE BCU0205B gene #SEQ ID 81.

xx	Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW	drug discovery; clinical medicine; forensic medicine; gene;
KW	chromosome 5q33.3; ds.

Homo sapiens.

WO2003029421-A2.

10-APR-2003.

02-OCT-2002: 2002WO-US031287.

03-OCT-2001. 2001UIS-0326526P.

14-MAY-2002; 2002US-00144194.

(ORIG-) ORIGENE TECHNOLOGIES INC.

Sun Z., Li X., Fan W., Kovacs KF, Jay G:

WPT: 2003-381623/36.

P-PSDB; ABR58318.

New isolated human differentially-regulated breast cancer polynucleotide and polypeptide, useful for diagnosing, staging, prognosticating, preventing and/or treating diseases and conditions relating to breast cancer.

Claim 2: SEO ID NO 81: 127pp + Sequence Listing; English.

The invention relates to isolated polynucleotides which are differentially-regulated in breast cancer. The methods and compositions of the present invention are useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing and/or treating diseases and conditions relating to breast cancer, and may be used in gene therapy or antisense therapy. They can also be used in research, drug discovery, clinical medicine and forensic medicine. Sequences given in records ACC72012-ACC72074 represent polynucleotides of the invention that are differentially-regulated in breast cancer. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct/sequences](http://wipo.int/pub/published/pct/sequences)

Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other; XX

Alignment Scores:		
Pred. No.:	8.64e-51	9695
Score:	948.00	207
Percent Similarity:	64.06%	Conservative: 55
Best Local Similarity:	50.61%	Mismatches: 91
Query Match:	44.24%	Indels: 56
DB:	7	Gaps: 13

US-10-029-020-14 COPY 1 400 (1-400) x ACC72052 (1-9695)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19

Db 435 ATGGATGTAAGGACCGCGA--CACCGCTCTTTGACCAGGACGCTGTGGCAAAGAG 491

Qv 20 ArgArgTyrThrSerSerSerAlaAspSerGluGlyLysAlaPro---GlnLysSer 38

492 TGTGGCTACACAAGCTCCTCTCTGGACAGTGAGGACTGCCGGTGCCACACAGAAATCC 551

Db 1494 GGAAACGATGATGGCAACAATGCCA 1520
RESULT 13
AASI4089
ID AAS14089 standard; DNA; 9729 BP.
XX AAS14089;
AC
XX 18-DEC-2001 (first entry)
XX
XX Human FCTR3f DNA sequence.
DE
XX Human; FCTRX; myelogenous leukaemia; carcinoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greenouw type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
XX
OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT 5'UTR 1..209
FT /*tag= b
FT CDS 210..8384
FT /*tag= a
FT /*product= "Human FCTR3f"
FT 3'UTR 8385..9729
FT /*tag= c
XX
WO200166747-A2.
PN
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US007160.
XX
PR 03-MAR-2000; 2000US-0186592P.
PR 03-MAR-2000; 2000US-0186718P.
PR 06-MAR-2000; 2000US-0187293P.
PR 06-MAR-2000; 2000US-0187294P.
PR 17-MAR-2000; 2000US-0190400P.
PR 07-APR-2000; 2000US-0196018P.
PR 03-JAN-2001; 2001US-0259548P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2001-596837/67.
XX P-PSDB; AAU08681.
XX
PT Novel polypeptides designated as FCTRX polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
PS Claim 9; Page 37-39; 215pp; English.
XX
XX The invention relates to human FCTRX polypeptides, FCTR1-FCTR7, and the
CC nucleic acids encoding them. These sequences are useful for the treatment
CC or prevention of numerous disorders including myelogenous leukaemia,
CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal

CC alloimmune thrombocytopaenia, neurological disorders, neurodegenerative
CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greenouw type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC the FCTR3a homologue FCTR3f
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.21e-44 Length: 9729
Score: 839.00 Matches: 195
Percent Similarity: 54.42% Conservative: 51
Best Local Similarity: 43.14% Mismatches: 76
Query Match: 39.15% Indels: 130
DB: 5 Gaps: 14

US-10-029-020-14_COPY_1_400 (1-400) x AAS14089 (1-9729)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg--AspAlaGlu 19
Db 210 ATGGATGTAAGGACCGGCGA---CACCGCTCTTTGACCAGAGACGCTGTGGCAAAGAG 266
QY 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 267 TGTGCTACACAAGCTCCTCTCTGGACAGTGAGGACTGCCGGGTGCCACAGAAATCC 326
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 327 TACAGCTCCAGTGAGACTCTGAGGCCTATGACCATGACAGCAGGATGCATATGGAAAC 386
QY 59 ArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
Db 387 CGAGTCACAGACCTCATCCACCGGAGTCAGATGAGTTTCTTAGACAAGGAACCAACTTC 446
QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThr 98
Db 447 ACCCTTGCCGAACCTGGGCATCTGTGAGCCCTCC---CCACACCGAAGCGGTACTGTCTCC 503
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 504 GACATGGGGATCCTTACCAGGGCTACTCCTTAGACAGGGTCTGACGCCGACTCCGAC 563
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
Db 564 ACCGAGGGAGGGATGTCTCCAGAACACGCCCATCAGACTGTGGGGCAGAGGGATAAATCC 623
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 624 AGCGGCAGTTCGGGCCCTGTCAGATCGTGAAAACTCGGCCCTTACCCCTGACTGACTCTGAC 683
QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
Db 684 AACGAAAACAAATCAGATGATGAGAACGGTCGTCCCATTCACCTACATCCTCGCCTAGT 743
QY 168 ----- 168
Db 744 CTCCTCCCATCTGCTCAGTGCCTAGTCCCATATAATCCTCCACCAGTTAGTGCAGATG 803
QY 168 ----- 168
Db 804 CCATTGCTAGACAGCAACACCTCCCATCAAATCATGAGACACCAACCCTGATGAGGAATTC 863
QY 169 -----Gly 169
Db 864 TCCCCCAATTCACTCTGCTCAGAGCATGCTCAGGGCCCCCAGCAAGCCTCCAGCAGTGGC 923

Qy	170	LeuGlnAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla	186	XX	08-MAY-2003.
Db	924	CTCCGAACACACAGCCAGTCGACTCTGAGGCCCTCTCCACCC---CCTCACAAC	980	XX	05-MAR-2001; 2001US-00800198.
Qy	187	HisThrProAsnGlnHisHisAlaAaSerIleAsnSerLeuAsnArgGlyAsnPheThr	206	PR	03-MAR-2000; 2000US-0186592P.
Db	981	CACACGCTGTCCCATCACCCAC---TCGTCCGCCAACTCCCTCAACAGGAACCTCACTGACC	1037	XX	(VERM/) VERNET C.
Qy	207	ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly	222	PA	(FERN/) FERNANDES E.
Db	1038	AATCGCGGAGTCAGATCCACGCCCGCGCCCGCCAGCGCCCAATGAC-----CTGGCCACC	1091	PA	(SHIM/) SHIMKETS R.
Qy	223	GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsn	242	PA	(HERR/) HERRMANN J.
Db	1092	ACACCA-----GAGTCCGTTTCAGCTTCAGACAGCTGGGTGCTAAAC	1133	PA	(MAJU/) MAJUMDER K.
Qy	243	SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln	262	PA	(MACD/) MACDOUGALL J.
Db	1134	AGCAACGTGCCACTGGAGACCCGG-----	1157	PA	(MISH/) MISHRA V.
Qy	263	AspAsnIleLeuGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp	282	PA	(MEZE/) MEZES P S.
Db	1157	-----GTGCGCTGTCGTTGAAA	1412	PA	(RAST/) RASTELLI L.
Qy	283	GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro	301	XX	Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
Db	1158	---CACTTCCTCTTCAAGACCTCCTCGGGGAGCACACCCCTTGTTCAGCAGCTCTTCCCG	1214	PI	Macdougall J, Mishra V, Mezes PS, Rastelli L;
Qy	302	GlyTyrProLeuThrSerSerThrValTyrSerProProProProProLeuProArgSer	321	XX	WPI; 2003-625633/59.
Db	1215	GGATACCTTTGACCTCAGGAACGGTTTACACGCCCGCCCGCGCTGCTGCCAGGAAT	1274	DR	P-PSDB; ADB32029.
Qy	322	ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys	341	XX	New FCTRX polypeptide and encoding polynucleotide, useful for preventing
Db	1275	ACTTTTCCAGGAAGGCTTTCAGCTGAAGAAGCCCTCCAAATATCTGAGCTGGAATGT	1334	PT	or treating FCTRX-related disorders, such as cancer, autoimmune,
Qy	342	AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal	361	PT	neurodegenerative, gastrointestinal, reproductive and inflammatory
Db	1335	GCTGCCCTCTCGCCATTGCGCGGCCCTCTCTTGGCTATTTTGTGGGTATTTCATA	1394	XX	diseases.
Qy	362	AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln	373	PS	Claim 9; Page 34-37; 155pp; English.
Db	1395	-----GTGCGCTGTCGTTGAAA	1412	XX	The invention relates to FCTRX polypeptides and the polynucleotides
RESULT 14				CC	encoding them. The sequences of the invention are useful for the
ADB32028				CC	manufacture of a medicament for diagnosing and treating disorders
ID	ADB32028	standard; cDNA; 9729 BP.		CC	associated with the FCTRX polypeptide, such as colorectal cancer,
XX				CC	adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
AC	ADB32028;			CC	autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
XX				CC	tumours, mammary tumours, human gliomas, astrocytomas, renal cell
DT	04-DEC-2003 (first entry)			CC	carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
XX				CC	cell and granular cell carcinomas, neurological disorders,
DE	Human FCTR3f cDNA.			CC	neurodegenerative disorders, nerve trauma, familial myelodysplastic
XX				CC	syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
KW	Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;			CC	conditions, immunological disorders, allergy and infection, asthma, lung
KW	myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;			CC	diseases, male and female reproductive disorders, deafness, glycoprotein
KW	ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;			CC	deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
KW	renal cell carcinoma; melanoma; clear cell carcinoma;			CC	C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
KW	granular cell carcinoma; neurological disorder;			CC	infection, spinocerebellar ataxia, plasmodium falciparum infection,
KW	neurodegenerative disorder; nerve trauma;			CC	Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
KW	familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;			CC	represents cDNA encoding an FCTRX polypeptide of the invention.
KW	Gardner syndrome; mental health condition; immunological disorder;			XX	Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
KW	allergy; asthma; lung disease; reproductive disorder; deafness;			SQ	
KW	glycoprotein deficiency; desmoid tumour; turcot syndrome;				
KW	liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;				
KW	diabetes; schistosoma mansoni infection; spinocerebellar ataxia;				
KW	plasmodium falciparum infection; Groenouw's corneal dystrophy;				
KW	lattice corneal dystrophy.				
XX					
OS	Homo sapiens.				
XX					
PN	US2003087816-A1.				

QY 59 ArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
Db 387 CGAGTCACAGACCTCATCCACCGGGAGTCAGATGATTTCTCTAGACAGGAACCAACTTC 446
QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThr 98
Db 447 ACCCTTGCGAAGTGGCATCTGTGAGCCCTCC---CCACACCGAAGCGGCTACTGCTCC 503
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 504 GACATGGGATCCTTCACAGGGCTACTCCCTTAGCACAGGGTCTGACGCCGACTCCGAC 563
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer 138
Db 564 ACCGAGGAGGATGTCTCCAGAACACGCCATCAGACTGTGGGCAGAGGGATAAAATCC 623
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 624 AGGCGAGTTCGGCCTGTCCAGTCGTGAAGAACTCGGCCCTTACCTGACTGACTCTGAC 683
QY 159 HisGluAsnThrGluThrAspHisProGly-----168
Db 684 AACGAAACAAATCAGATGATGAGAACGGTCGTCCCATTCACCTACATCCTCGCCTAGT 743
QY 168 -----168
Db 744 CTCCTCCCATCTGCTCAGCTGCCTAGCTCCCATATCCTCCACAGTTAGTGCCAGATG 803
QY 168 -----168
Db 804 CCATTGCTAGACAGCAACACCTCCCATCAATCATGGACACCAACCCTGATGAGGAATTC 863
QY 169 -----Gly 169
Db 864 TCCCCCAATTATACCTGCTCAGAGCATGCTCAGGGCCCCAGCAAGCCTCCAGCAGTGGC 923
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProLeuSerHisAla 186
Db 924 CCTCGGAACCAACACAGCCAGTCGACTCTGAGGCCCCCTCTCCACCC---CCTCACAAAC 980
QY 187 HisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
Db 981 CACACGCTGTCCCATCACAC---TCGTCCGCCAACTCCCTCAACAGGAACCTCACTGACC 1037
QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
Db 1038 AATCGGCGGAGTCAGATCCACGCCCGCGGCCAGCGCCCAATGAC-----CTGGCCACC 1091
QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsn 242
Db 1092 ACACCA-----GAGTCCGTTCAGCTTCAGGACAGCTGGGTGCTAAAC 1133
QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
Db 1134 AGCAACGTGCCACTGGAGACCCGG-----1157
QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
Db 1157 -----1157
QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
Db 1158 ---CACTTCCTCTTCAAGACCTCCTCGGGAGCACACCCCTGTTCAGCAGCTCTTCCCCG 1214
QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArgSer 321
Db 1215 GGATACCCCTTTGACCTCAGGAACGGTTTACACGCCCCCCCGCCCTGCTGCCCAAGAT 1274
QY 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys 341
Db 1275 ACTTCTCCAGGAAGGCTTCAAGCTGAAGAAGCCCTCCAAATACTGCAGCTGGAATGT 1334
QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361

Db 1335 GCTGCCCTCTCCGCCATTGCCGCGGCCTCCTCTTGCTATTTTGGCTATTTCGGCTATTTCATA 1394
QY 362 AlaMethHisLeuPheGlyLeuAsnTrpHisLeuGln 373
Db 1395 -----GTGCCCTGGTCGTTGAAA 1412
RESULT 15
AAS14085
ID AAS14085 standard; DNA; 9826 BP.
XX
AC AAS14085;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human FCTR3b DNA sequence.
XX
KW Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greonow type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW antiallergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3b; neurestin-like protein.
XX
OS Homo sapiens.
XX
FH
FT 5'UTR Location/Qualifiers
FT 1..279 /*tag= b
FT CDS 280..8481 /*tag= a
FT /*product= "Human FCTR3b"
FT 3'UTR 8482..9826 /*tag= c
FT
PN WO200166747-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US007160.
XX
PR 03-MAR-2000; 2000US-0186592P.
PR 03-MAR-2000; 2000US-0186718P.
PR 06-MAR-2000; 2000US-0187293P.
PR 06-MAR-2000; 2000US-0187294P.
PR 17-MAR-2000; 2000US-0190400P.
PR 07-APR-2000; 2000US-0196018P.
PR 03-JAN-2001; 2001US-0259548P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2001-596837/67.
DR P-PSDB; AAU08680.
XX
PT Novel polypeptides designated as FCTRX polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
PS Claim 9; Page 33-35; 215pp; English.
XX

CC The invention relates to human FCTR polypeptides, FCTR1-FCTR7, and the
CC nucleic acids encoding them. These sequences are useful for the treatment
CC or prevention of numerous disorders including myelogenous leukaemia,
CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
CC alloimmune thrombocytopenia, neurological disorders, neurodegenerative
CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greenuw type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers Corneal dystrophy. This sequence represents DNA encoding
CC FCTR3b, a neurestin-like protein
XX
SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;

Alignment Scores: Length: 9826
Pred. No.: 8.3e-44 Matches: 195
Score: 839.00
Percent Similarity: 54.42% Conservative: 51
Best Local Similarity: 43.14% Mismatches: 76
Query Match: 39.15% Indels: 130
DB: 5 Gaps: 14

US-10-029-020-14_COPY_1_400 (1-400) x AAS14085 (1-9826)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
DB 280 ATGGATGTAAGGACCGCGA---CACCGCTCTTTGACCAGAGGACGCTGTGGCAAAGAG 336
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
DB 337 TGTCGCTAGACAAAGCTCCTCTCTGGACAGTGAGGACTGCCGGGTGCCACACAGAAATCC 396
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
DB 397 TACAGCTCCAGTGAGACTCTGAAGGCCTATGACCATGACAGCAGGATGCACTATGGAAC 456
QY 59 ArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
DB 457 CGAGTCACAGACCTCATCCACCGGAGTGAGATGAGTTTCTTAGACAGGAACCAACTTC 516
QY 79 ThrLeuArgGluLeuGluGluGluValThrProProHisGlyThrLeuTyrArgThr 98
DB 517 ACCCTTGCCGAACCTGGGATCTGTGAGCCCTCC---CCACACCGAAGCGGCTACTGCTCC 573
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
DB 574 GACATGGGATCCTTCACAGGGCTACTCCCTTAGCACAGGGTCTGACCGGACTCCGAC 633
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrPheGlyArgSerThrArgSer 138
DB 634 ACCGAGGAGGATGTCTCCAGAACACGCCATCAGACTGTGGGGCAGAGGGATAAATCC 693
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
DB 694 AGGCGCAGTTCGGCCTGTCCAGTCGTGAAACTCGGCCCTTACCCCTGACTGACTCTGAC 753
QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
DB 754 AACGAAACAAATCAGATGATGAGAACGGTGGTCCCATTCACCTACATCCTCGCCTAGT 813
QY 168 ----- 168
DB 814 CTCCTCCCATCTGCTCAGCTGCCTAGTCCCATATAATCCTCCACCAAGTTAGTGCCAGATG 873
QY 168 ----- 168
DB 874 CCATTGCTAGACAGCAACACCTCCCATCAATCATGGACACCAACCCCTGATGAGGAATTC 933

QY 169 -----Gly 169
DB 934 TCCCCAATTACCTGCTCAGAGCATGCTCAGGGCCCCCAGCAAGCCTCCAGCAGTGGC 993
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla 186
DB 994 CTTCCGAACCAACACAGCCAGTCGACTCTGAGGCCCTCTCCACCC---CCTCACAAC 1050
QY 187 HisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
DB 1051 CACAGCTGTCCCATCACCAC---TCGTCCGCCAACTCCCTCAACAGGAACCTACTGACC 1107
QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
DB 1108 AATCGCGGAGTTCAGATCCACGCCCGCCGCGCCAGCGCCCAATGAC-----CTGGCCACC 1161
QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTyrLeuLeuAsn 242
DB 1162 ACACCA-----GAGTCCGTTTCAGCTTCAGGACAGCTGGGTGCTAAAC 1203
QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
DB 1204 AGCAACGTGCCACTGGAGACCCGG----- 1227
QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
DB 1227 ----- 1227
QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
DB 1228 ---CACTTCCTCTTCAAGACCTCTCTCGGGGAGCACACCTTTGTTACAGCAGCTCTTCCCCG 1284
QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProProArgProLeuProArgSer 321
DB 1285 GGATACCTTTGACCTCAGGAACGGTTTACACGCCCGCCCGCCGCTGCTGCCAGGAAT 1344
QY 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTyrLysCys 341
DB 1345 ACTTTCTCAGGAAGGCTTTCAAGCTGAAGAAGCCCTCCAAATACAGCTGGAATGT 1404
QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPheVal 361
DB 1405 GCTGCCCTCTCCGCCATTGCGCGGCCCTCTCTTGGCTATTTTGTGGCGTATTTCATA 1464
QY 362 AlaMetHisLeuPheGlyLeuAsnTyrHisLeuGln 373
DB 1465 -----GTGCCCTGGTGGTGA 1482

Search completed: August 14, 2004, 02:28:40
Job time : 745.508 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 801.682 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400
Perfect score: 2143
Sequence: 1 MDVKRKPYSLTRRDAER.....EITEDTASSWPVPTDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published_Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020_@CGN_1_1_2156_@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match Length	ID	Description
* US-10-383-201-43					

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

1	2143	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
2	2143	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
3	2138	99.8	1680	15	US-10-290-578-3	Sequence 3, Appli
4	2138	99.8	3111	9	US-09-773-517-12	Sequence 12, Appl
5	2138	99.8	3111	9	US-09-792-025-12	Sequence 12, Appl
6	2138	99.8	3111	9	US-09-849-868-12	Sequence 12, Appl
7	2138	99.8	3111	15	US-10-290-578-1	Sequence 1, Appli
8	2138	99.8	3111	15	US-10-453-183-12	Sequence 12, Appl
9	2130	99.4	8355	13	US-10-383-201-55	Sequence 55, Appl
10	2098	97.9	8438	13	US-10-042-865-1	Sequence 1, Appli
11	1285	60.0	2387	15	US-10-290-578-11	Sequence 11, Appl
12	987	46.1	8645	17	US-10-038-854-37	Sequence 37, Appl
13	987	46.1	8675	17	US-10-038-854-35	Sequence 35, Appl
14	948	44.2	9695	16	US-10-144-194A-81	Sequence 81, Appl
15	920	42.9	8689	9	US-09-808-602-78	Sequence 78, Appl
16	920	42.9	8689	10	US-09-800-198-66	Sequence 66, Appl
17	917	42.8	2496	9	US-09-808-602-75	Sequence 75, Appl
18	917	42.8	2496	10	US-09-800-198-63	Sequence 63, Appl
19	916	42.7	8797	9	US-09-808-602-74	Sequence 74, Appl
20	916	42.7	8797	9	US-09-808-602-77	Sequence 77, Appl
21	916	42.7	8797	10	US-09-800-198-62	Sequence 62, Appl
22	916	42.7	8797	10	US-09-800-198-65	Sequence 65, Appl
23	897	41.9	8409	9	US-09-808-602-79	Sequence 79, Appl
24	897	41.9	8409	10	US-09-800-198-67	Sequence 67, Appl
25	871	40.6	8575	13	US-10-072-012-143	Sequence 143, App
26	839	39.2	9729	9	US-09-808-602-12	Sequence 12, Appl
27	839	39.2	9729	10	US-09-800-198-12	Sequence 12, Appl
28	839	39.2	9826	9	US-09-808-602-7	Sequence 7, Appli
29	839	39.2	9826	10	US-09-800-198-7	Sequence 927, App
30	734.5	34.3	12880	16	US-10-295-027-927	Sequence 39, Appl
31	653.5	30.5	8473	17	US-10-038-854-39	Sequence 41, Appl
32	653.5	30.5	8487	17	US-10-038-854-41	Sequence 17, Appl
33	534	24.9	1755	10	US-09-998-966-17	Sequence 17, Appl
34	534	24.9	1755	15	US-10-004-415-17	Sequence 17, Appl
35	534	24.9	1755	16	US-10-384-974-17	Sequence 17, Appl
36	508	23.7	1727	16	US-10-094-749-710	Sequence 710, App
37	508	23.7	9058	16	US-10-144-194A-79	Sequence 79, Appl
38	454	21.2	527	15	US-10-029-386-1356	Sequence 1356, Ap
39	434	20.3	1430	9	US-09-808-602-5	Sequence 5, Appli
40	434	20.3	1430	10	US-09-800-198-5	Sequence 5, Appli
41	434	20.3	1431	10	US-09-977-418-21	Sequence 21, Appl
42	434	20.3	1431	10	US-09-977-033A-21	Sequence 21, Appl
43	434	20.3	1431	10	US-09-977-751C-21	Sequence 21, Appl
44	434	20.3	1431	10	US-09-977-639A-21	Sequence 21, Appl
45	434	20.3	1431	11	US-09-977-819B-21	Sequence 21, Appl


```
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 4.85e-207 Length: 8354
Score: 2143.00 Matches: 400
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-10-383-201-43 (1-8354)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 35 ATGGACGTGAAGAGAGAGGAGCCCTTACCGCTCGCTGACCCCGCGCGACGCGGAGCGC 94
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 95 CGGTACACCAAGCTGCTCCGGGACAGCGAGGAGGCAAGCCCGCGAGAAATCGTACAGC 154
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 155 TCCAGCGAGACCTGAAGGCCTACGACCAGAGCGCCCGCCTAGCCTATGGCAGCGCGGTC 214
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 215 AAGGACATTGTGCCGAGGAGGCGGAGGAATCTGCCGCACACAGGTGCCAACTTCACCCGTG 274
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 275 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGACAGACATT 334
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 335 GGCCTGCCCAATGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGCTGAC 394
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 395 ACGGTGCTGTCCCTTGAGCACCCCGTGTGTCGTGTCGGGGCGGAGCACACGGTCAGGGCGC 454
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 455 AGCTCTGCCTGTCCAGCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 514
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 515 AACACTGAGACTGATCATCGGGCGGCTGCAGAACCAACGCGGCTCCGAGCGCGCGG 574
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 575 CCGCGGCTCTCGCACGCGCCACACCCCCAACCCAGCACCAACGCGGCTCCATTAACTCCCTG 634
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 635 AACCGGGGCAACTTCACGCGGAGGAGCAACCCAGCGCGGCCCCCAGCGGACCACTCGCTC 694
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
```

```
Db 695 TCCGAGAGCCCCCTGCCGGGCGGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG 754
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 755 CTCACAGCAACATCCCCCTGGAGACCAGGAACCTAGGAAGACGACCATTCCTAGGGACA 814
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 815 TTGCAGGACAACCTCATTTGAGATGGACATTTCTCGGCGCCTCCCGCATGATGGGGCTTAC 874
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 875 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCATCA 934
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 935 CCAGGTACCCACTGACGTCCAGCACAGTGACTCTCTCCGCCCCGACCCCTGCCCCGC 994
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 995 AGCACCTTCGCCCGCGGCCCTTTAACCTCAAGAAGCCCTCCAACTACTGTAACCTGGAAG 1054
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1055 TGCGCAGCCCTGAGCGCCATCGTCACTCTCAGCCACTCTGGTCACTCTGCTGGCATACTTT 1114
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1115 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1174
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1175 GAGATCACGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCGACGTCCTCCCTATACCCC 1234

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.: 4.85e-207 Length: 8354
```

Score:		2143.00	Matches:	400
Percent Similarity:		100.00%	Conservative:	0
Best Local Similarity:		100.00%	Mismatches:	0
Query Match:		100.00%	Indels:	0
DB:		13	Gaps:	0
US-10-029-020-14_COPY_1_400 (1-400) x US-10-029-020-13 (1-8354)				
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20	
Db	35	ATGGACGTGAAGGAGAGAACCTTACCGCTCGCTGACCCGGCGCGACGCCGAGCGC	94	
QY	21	ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40	
Db	95	CGCTACACAGCTCGTCCGCGACAGCGAGGAGGGCAAGCCCGCAGAAATCGTACAGC	154	
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60	
Db	155	TCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGCGTC	214	
QY	61	LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu	80	
Db	215	AAGGACATTGTGCCGCGAGGAGGCCGAGGAATTCTGCCGCACACGGTGCCAATTCAACCTG	274	
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100	
Db	275	CGGGAGCTGGGGCTGGAAGAGTAACGCCCCCTCACGGGACCTGTACCGGACAGACATT	334	
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120	
Db	335	GGCCTGCCCCAAATGCGGCTACTCCATGGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	394	
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140	
Db	395	ACGGTGCTGTCCCTGAGCACCCCGTGGTCTGTGGGGCCGGAGCACACGGTCAGGGCGC	454	
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160	
Db	455	AGCTCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACGACACCGAGCATGAA	514	
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180	
Db	515	AACACTGAGACTGATCATCCGGGGCGCTGCAGAACCAACCGCGGGCTCCGGACGCCGCCG	574	
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200	
Db	575	CCGGCGCTCTGCACGCCCCACACCCCCAAACCAAGCACCAACCGCGGCTCCATTAACTCCCTG	634	
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220	
Db	635	AACGGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCACGGACCACTCGCTC	694	
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240	
Db	695	TCCGGAGAGCCCCCTGCCGGCGGGCCAGGAGCCTGCCACGCCCAGGAGAACTGGCTG	754	
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr	260	
Db	755	CTCAACAGCAACATCCCCCTGGAGACCAGGAACCTAGGCAAGCAGCCATTCTTAGGGACA	814	
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280	
Db	815	TTGCAGGACAACCTCATTTGATGTGACATTCTCGGCGCCTCCCGCCATGATGGGGCTTAC	874	
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300	
Db	875	AGTGACGGCACTTCTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCAACATCA	934	
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg	320	
Db	935	CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGGCCCGGACCCCTGCCCGCG	994	
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340	

Db	995	AGCACCTTCGCCCGCGCCCTTTAACTCAAGAAGCCCTCAAGTACTGTAAGTGAAG	1054	
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe	360	
Db	1055	TGCGCAGCCCTGAGCGCCATCGTCACTTCAGCCACTCTGGTCACTCCTGCTGGCATACTTT	1114	
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380	
Db	1115	GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT	1174	
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400	
Db	1175	GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC	1234	
RESULT 3				
US-10-290-578-3				
; Sequence 3, Application US/10290578				
; Publication No. US20030078389A1				
; GENERAL INFORMATION:				
; APPLICANT: Schaefer, Gabriele M.				
; Sliwkowski, Mark				
; TITLE OF INVENTION: Gamma-Herregulin				
; NUMBER OF SEQUENCES: 11				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Genentech, Inc.				
; STREET: 460 Point San Bruno Blvd				
; CITY: South San Francisco				
; STATE: California				
; COUNTRY: USA				
; ZIP: 94080				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: WinPatin (Genentech)				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/10/290,578				
; FILING DATE: 08-No. US20030078389A1-2002				
; CLASSIFICATION: <Unknown>				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: US/08/891,845				
; FILING DATE: <Unknown>				
; APPLICATION NUMBER: 60/021640				
; FILING DATE: 12-Jul-96				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Lee, Wendy M.				
; REGISTRATION NUMBER: 40,378				
; REFERENCE/DOCKET NUMBER: P1043				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 415/225-1994				
; TELEFAX: 415/952-9881				
; TELEX: 910/371-7168				
; INFORMATION FOR SEQ ID NO: 3:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 1680 base pairs				
; TYPE: Nucleic Acid				
; STRANDEDNESS: Single				
; TOPOLOGY: Linear				
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:				
US-10-290-578-3				
Alignment Scores:				
Pred. No.:		1.84e-207	Length:	1680
Score:		2138.00	Matches:	399
Percent Similarity:		99.75%	Conservative:	0
Best Local Similarity:		99.75%	Mismatches:	1
Query Match:		99.77%	Indels:	0
DB:		15	Gaps:	0
US-10-029-020-14_COPY_1_400 (1-400) x US-10-290-578-3 (1-1680)				
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20	

Db 1 ATGGACGTGAAGGAGAGGAAGCCCTTACCGCTCGTGACCCCGCGCGGACGCCGAGCGC 60
Qy 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 61 CGCTACACCACTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC 120
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 121 TCCAGCGAGACCCCTGAAGGCCTACGACAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC 180
Qy 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 181 AAGGACATTGTGCCGAGGCGGAGGCGGAGGAATTCTGCCGACACAGGTGCCAACTTCACCCGTG 240
Qy 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 241 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 300
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaaspMetGluAlaAsp 120
Db 301 GGCCTCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 360
Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 361 ACGGTGCTGCCCTTGAGCACCCCGTGCTGTGGGGCCGGAGCACACGGTCAAGGGCGC 420
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 421 AGCTCCTCGCTGCCAGCGGGCCAAATTCCAATCTCACACTCACCGACACCGAGCATGAA 480
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 481 AACACTGAGACTGATCATCCGGCGGCTGCAGAACCAACGCGCGGCTCCGGACGCCGCG 540
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 541 CCGCCGCTCTGCACGCCCCACACCCCCAACAGCACCAACGCGGCTCCATTAACTCCCTG 600
Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 601 AACCGGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCCGACCGACCACTCGCTC 660
Qy 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 661 TCCGAGAGCCCTGCGCGGCGGCCAGAGACCTGCCACGCCCCAGGAGAACTGGCTG 720
Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr 260
Db 721 CTCAACAGCAACATCCCTCTGGAGACCAAGAACCTAGGCAAGCCATTCCTAGGGACA 780
Qy 261 LeuGlnAspAsnIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 781 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGCTTAC 840
Qy 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 841 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCACATCA 900
Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 901 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGCG 960
Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 961 AGCACCTTCGCGCGCGCGGCTTTAACTTCAAGAAGCCCTCCAAGTACTGTAACTGGGAG 1020
Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1021 TCGCAGCCCTGAGCGCCATCGTCTCTCAGCCACTCTGGTCTATCTCTGTCGCATACTTT 1080
Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1081 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1140
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1141 GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCCCTATACCCC 1200
RESULT 4
US-09-773-517-12
; Sequence 12, Application US/09773517
; Patent No. US20010023241A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/09/773,517
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/243,198
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-773-517-12
Alignment Scores:
Pred. No.: 4,19e-207 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x US-09-773-517-12 (1-3111)
Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGGAGAGGAAGCCCTTACCGCTCGTGACCCGGCGCGCGAGCGC 393
Qy 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCACTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC 453
Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGGCCTACGACAGGACGCCCGCTAGCCTATGGCAGCCGCGTC 513
Qy 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTGTCCCGCAGGAGGCGCGGAGGAATTCTGCCGACAGGTGCCAACTTCACCCGTG 573
Qy 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGACAGACATT 633
Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaaspMetGluAlaAsp 120
Db 634 GGCCTCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGCCCTGAGCACCCCGTGCTGTGGGGCCGGAGCACACGGTCAAGGGCGC 753
Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCCTGCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCGCTTCAGAAACCAACGCGCGGCTCCGGACGCCGCG 873
Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200

Db 874 CCGCCGCTCTCGCAGCGCCACACCCCAACGACGACGCGGCTCCATTAACTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220

Db 934 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCGGGCCCCACGACCACTCGCTC 993

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240

Db 994 TCCGAGAGCCCCCTGCGGCGGCGCCAGGAGCTGCCACGCGCCAGGAGAACTGGCTG 1053

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260

Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAGAACTAGGCAAGCAGCCATTCTTAGGGACA 1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280

Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGGGCTCCCGCCATGATGGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300

Db 1174 AGTGACGGGCACTTCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCAATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320

Db 1234 CCAGGGTACCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCTGCCCCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340

Db 1294 AGCACCTTCGCGCGCGCGGCTTTAACTCAAGAAGCCCTCCAAAGTACTGTAAGTGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360

Db 1354 TGCGCAGCCCTGAGCGCCATCGTCTATCTCAGCCACTCTGGTCTATCTGTGGCATACTT 1413

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1414 GTGGCCATGCACCTGTTTGGCCATAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1474 GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCGACGTCTCCCTATACCC 1533

RESULT 5

US-09-792-025-12

Sequence 12, Application US/09792025

Patent No. US20020042087A1

GENERAL INFORMATION:

APPLICANT: Sliwowski, Mark

Kern, Jeffrey

TITLE OF INVENTION: Use of Heregulin as a Growth Factor

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/792,025

FILING DATE: 23-Feb-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/020,598

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Schwartz, Timothy R.

REGISTRATION NUMBER: 32171

REFERENCE/DOCKET NUMBER: P1145

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-7467

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 3111 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

MOLECULE TYPE: nucleic acid

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-792-025-12

Alignment Scores:

Pred. No.: 4.19e-207 Length: 3111

Score: 2138.00 Matches: 399

Percent Similarity: 99.75% Conservative: 0

Best Local Similarity: 99.75% Mismatches: 1

Query Match: 99.77% Indels: 0

DB: 9 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-09-792-025-12 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20

Db 334 ATGGACGTGAAGGAGAGGAGAGCCTTACCGCTCGCTGACCCGCGCGCGAGCGCGC 393

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40

Db 394 CGCTACACCACTCGTCCCGGACAGCGAGGAGGCAAGCCCGCGCTAGCTATGGCAGCGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60

Db 454 TCCAGCGAGACCTTGAAGGCCTACGACCAAGACGCGCGCTAGCTATGGCAGCGCGC 513

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80

Db 514 AAGGACATTGTCCCGCAGGAGGCGGAGGAATTCTCCCGCACAGGTGCCAACTTCACTC 573

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100

Db 574 CCGGAGCTGGGGCTGGAAGAAGTAACGCCCTTACCGGACCTGTACCGACAGACATT 633

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120

Db 634 GGCTTCCCCCACTGCGGCTACTTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693

QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140

Db 694 ACGTGCTGTCCCCCTGAGCACCCCTGCTGTGTGGGCGGAGCACACGGTTCAGGCGCG 753

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160

Db 754 AGCTCTGCCTGTCCAGCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180

Db 814 AACACTGAGACTGATCATCCGGGCGGCTGCAGAACCCACGCGGGCTCCGAGCGCGCG 873

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200

Db 874 CCGCCGCTCTCGCACGCCACACCCCCAACCCAGCACCCAGCGCGCTCCATTACTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220

Db 934 AACCGGGCAACTTCACGCGCGGAGGAGCAACCCAGCGCGCGCCCCACGACCACTCGCTC 993

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240

Db 994 TCCGAGAGCCCCCTGCGGCGGCGCGCGAGGAGCTGCCACCGCCAGGAGAACTGGCTG 1053

QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
DB	1054	CTCAACAGCAACATCCCTCGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA	1113
QY	261	LeuGlnAspAsnIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
DB	1114	TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGCTTAC	1173
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
DB	1174	AGTGACGGGCACITCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCACATCA	1233
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg	320
DB	1234	CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCTGCCCGC	1293
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys	340
DB	1294	AGCACCTTCGCCCGCGCGCCCTTTAACCTCAAGAACCCCTCCAAGTACTGTAACTGGAAG	1353
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe	360
DB	1354	TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCACTCTGTGGCATACTTT	1413
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
DB	1414	GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAAGCCGATGGAGGGGCAGATGTAT	1473
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
DB	1474	GAGATCACGGAGGACACAGGCGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC	1533

RESULT 6

```

US-09-849-868-12
; Sequence 12, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Gao, Wei-Qiang
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)... (2637)
US-09-849-868-12

```

Alignment Scores:		
Pred. NO.:	4.19e-207	3111
Score:	2138.00	399
Percent Similarity:	99.75%	Conservative: 0
Best Local Similarity:	99.75%	Mismatches: 1
Query Match:	99.77%	Indels: 0
DB:	9	Gaps: 0

U.S.-10-029-020-14 COPY 1 400 (1-400) x US-09-849-868-12 (1-3111)

QY	1	Met	Asp	Val	Lys	Glu	Arg	Lys	Pro	Tyr	Arg	Ser	Leu	Thr	Arg	Arg	Asp	Ala	Glu	Arg	20	
DB	334	ATG	GAC	GTG	AAG	GAG	AGC	CTT	ACC	GCT	CGC	TGAC	CCG	CGC	CGC	GCG	AGC	GCG	CGC	393		
QY	21	Arg	Tyr	Thr	Ser	Ser	Ala	Asp	Ser	Glu	Glu	Gly	Lys	Ala	Pro	Gln	Lys	Ser	Tyr	Ser	40	
DB	394	CGT	TAC	CAC	CAG	CTCG	TCC	GCG	GAC	AGC	GAG	GAG	CAA	AAG	CCC	CGC	GAAA	TCT	GTA	CAG	453	
QY	41	Ser	Ser	Glu	Thr	Leu	Lys	Ala	Tyr	Asp	Gln	Asp	Ala	Arg	Glu	Ala	Tyr	Gly	Ser	Arg	Val	60

RESULT 7

US-10-290-578-1

; Sequence 1, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Hereregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-No. US20030078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-290-578-1

Alignment Scores:
Pred. No.: 4.19e-207 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 15 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x US-10-290-578-1 (1-3111)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGGAGGAAGCCCTTACCGCTCGCTGACCCGGCGCGAGCGAGCGC 393
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACACAGCTCGTCGGGACAGGAGGGGCAAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGGCTACGACGACGAGCGCCCGCTAGCCTATGGCAGCGCGTC 513
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGACATTTGTCCGCGAGGCGGAGGAATTCTGCCGACAGGTGCCAACTTACCGCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100

Db 574 CGGAGCTGGGGCTGGAGAAGTAACGCCCTCTACGGGACCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 694 ACGTGCTGTCCCTGAGCACCCCGTGCCTCTGTGGGGCGGAGCACACGGTCAAGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCTGTCCAGCCGGGCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGCGGCTGCAGAACCAACCGCGGCTCCGGACGCCCGC 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTCGCACGCCACACCCCAACAGCACACCGCGGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTTCACGCCGAGGAGCAACCCAGCCCGCCCCACGAGCACCTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTyrLeu 240
Db 994 TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAAGAAACCTAGGCAAGCAGCATTCCTAGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACCTCATTGAGATGGACATTCCTCGGCGCTCCCGCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCTCTTCAAGCCTTGAGGACACCTCCCGCTCTTCTGCACCATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTACCCACTGACGTCCAGCACAGTACTCTCTCCGCCCGGACCCCTGCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTyrLys 340
Db 1294 AGCACCTTCGCCGGCGCGGCTTTAAACCTCAAGAGCCCTCCAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1354 TGCGCAGCCCTGAGCGCCATCGTCTCATCTCAGCCACTCTGGTCTCTCTGGCATACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCACGGAGGACACAGCCAGCAGTTGGSCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533
RESULT 8
US-10-453-183-12
; Sequence 12, Application US/10453183
; Publication No. US2003019429A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/10/453,183
; CURRENT FILING DATE: 2003-06-03

; PRIOR APPLICATION NUMBER: US/09/243,198
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,866
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-453-183-12

Alignment Scores:
Pred. No.: 4.19e-207 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 15 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-10-453-183-12 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGAGAGGAAGCCTTACCGCTCGCTGACCGCGCGCGCGAGCGC 393

QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACACAGCTCTCCGGGACAGCGAGGAGGSCAAAGCCCCCGCAGAAATCGTACAGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCTTGAAGGCCTACGACCGAGCGCCCGCTAGCCTATGGCAGCCGCGTC 513

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTGTGCGCAGGAGGCGCGAGGAATTCTGCCGCACAGGTGCCAACTTCAACCCTG 573

QY 81 ArgGluLeuGlyGluGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCTGTACCGGACAGACATT 633

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyValaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCTTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCCGTGGCTGTGGGGCCGGAGCACACGGTCAGGGCGC 753

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCCTGCCTGTCCAGCGCGGCCAATCCAAATCTCACACTCACCGACACCGAGCATGAA 813

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGCGGCTGCAGAACCAACCGCGGCTCCGAGCGCGCGC 873

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGGCTCTGCACGCCCCACACCCCAACACAGCACCGCGGCTCCATTAACCTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTACCGCGAGGAGCAACCCAGCCCGCGCCCGCCACGGACCACTCGCTC 993

QY 221 SerGlyGluProProAlaGlyGlyValaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGCCCCCTGCGGCGCGCCAGGAGCCTGCCACGCCCGAGGAACTGGCTG 1053

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTCTAGGACA 1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACCTCATTTAGATGGACATTCTCGGCGCTCCGCCATGATGGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 1234 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCCGCCCGACCCCTGCCCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGCGCCTTTAACTCAGAACCCCTCCAAGTACTGTAACTGGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCACTCCTGCTGCATACTTT 1413

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTGGCCTAACTGGCACCTGCAGCCGATGGAGGGGAGATGTAT 1473

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCACGGAGGACACAGCCAGCAGTTGGCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 9

US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 1.02e-205 Length: 8355
Score: 2130.00 Matches: 398
Percent Similarity: 99.50% Conservative: 0

Best Local Similarity: 99.50%				Mismatches: 2			
Query Match: 99.39%				Indels: 0			
DB: 13				Gaps: 0			
US-10-029-020-14_COPY_1_400 (1-400) x US-10-383-201-55 (1-8355)							
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20				
Db	1	ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCCGCGCGGACGCCGAGCGC	60				
QY	21	ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40				
Db	61	CGCTACACCACTGCTCCGCGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC	120				
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60				
Db	121	TCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCTATGGCAGCCCGCTC	180				
QY	61	LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu	80				
Db	181	AAGGACATTTGTGCCGAGGAGGCCGAGGAATTCTGCCGCACACAGGTGCCAACTTCACCCCTG	240				
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100				
Db	241	CGGGAGCTGGGGCTGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT	300				
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120				
Db	301	GGCCTCCCCCACTGGCGTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	360				
QY	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140				
Db	361	ACGGTGCTGTCCCTGAGCACCCCGTGCGTCTGTGGGGCCGAGCACACGGTCAGGGCGC	420				
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160				
Db	421	AGTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA	480				
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180				
Db	481	AACACTGAGACTGATCATCCGGCGGCTGCAGAACCCACGCGCGGCTCCGGACGCCGCCG	540				
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200				
Db	541	CGCGCGCTCTCGACGCCCCACACCCCCAACCCAGCACCCCGGCCCCACGGACCACTCGCTC	600				
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220				
Db	601	AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGGCCCCACGGACCACTCGCTC	660				
QY	221	SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240				
Db	661	TCCGGAGAGCCCCCTGCCGCGGCGGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG	720				
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260				
Db	721	CTCAACAGCAACATCCCTCGGAGACCAGAAACCTAGGCAAGCAGCCATTCTCTAGGGACA	780				
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280				
Db	781	TTGCAGGACAACCTCATTGAGATGGACATTCTCGGCGCTCCCGCCATGATGGGGCTTAC	840				
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300				
Db	841	AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACTCCCGCTCTTCTGCACCAATCA	900				
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320				
Db	901	CCAGGTATCCCACTGACGTCCAGCACAGTGATCTCTCTCCGCCCCGACCCCTGCCCGC	960				
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340				
Db	961	AGCACCTTCGCCTGGCCGGCCTTTAAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG	1020				

QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1021	TGCGCAGCCCTGAGCGCCATCGTCTCATCTCAGCCACTCTGGTCTATCTGCTGGCATACTTT	1080
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1081	GTGGCCATGCACCTGTTTGGCCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT	1140
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	1141	GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC	1200
RESULT 10			
US-10-042-865-1			
; Sequence 1, Application US/10042865			
; Publication No. US20040029216A1			
; GENERAL INFORMATION:			
; APPLICANT: Padigaru, Muralidhara			
; APPLICANT: Li, Li			
; APPLICANT: Zerhusen, Bryan D			
; APPLICANT: Casman, Stacie J			
; APPLICANT: Shenoy, Suresh G			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Zhong, Mei			
; APPLICANT: Gangolli, Esha A			
; APPLICANT: Burgess, Catherine E			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Vernet, Corine A.M			
; APPLICANT: Taylor, Sarah			
; APPLICANT: Tchernev, Velizar T			
; APPLICANT: Miller, Charles E			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Boldog, Ference L			
; APPLICANT: Grosse, William M			
; APPLICANT: Alsobrook II, John P			
; APPLICANT: Gerlach, Valerie L			
; APPLICANT: Edinger, Shlomit R			
; APPLICANT: Rothenberg, Mark E			
; APPLICANT: Ellerman, Karen			
; APPLICANT: MacDougall, John			
; APPLICANT: Malyankar, Uriel M			
; APPLICANT: Millet, Isabelle			
; APPLICANT: Peyman, John			
; APPLICANT: Smithson, Glennda			
; APPLICANT: Gunther, Erik			
; APPLICANT: Stone, David			
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of			
; TITLE OF INVENTION: Using the Same			
; FILE REFERENCE: 21402-537			
; CURRENT APPLICATION NUMBER: US/10/042,865			
; CURRENT FILING DATE: 2002-05-17			
; PRIOR APPLICATION NUMBER: 60/260,417			
; PRIOR FILING DATE: 2001-01-09			
; PRIOR APPLICATION NUMBER: 60/260,831			
; PRIOR FILING DATE: 2001-01-10			
; PRIOR APPLICATION NUMBER: 60/272,338			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: 60/274,876			
; PRIOR FILING DATE: 2001-03-09			
; PRIOR APPLICATION NUMBER: 60/284,704			
; PRIOR FILING DATE: 2001-04-18			
; NUMBER OF SEQ ID NOS: 264			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 8438			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-042-865-1			
Alignment Scores:			
Pred. No.: 1.86e-202			
Score: 2098.00			
Length: 8438			
Matches: 398			

Percent Similarity: 98.76%		Conservative: 0	
Best Local Similarity: 98.76%		Mismatches: 2	
Query Match: 97.90%		Indels: 3	
DB: 13		Gaps: 1	
US-10-029-020-14_COPY_1_400 (1-400) x US-10-042-865-1 (1-8438)			
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	4	ATGGACGTGAAGGAGAGAACTTACCGCTCGCTGACCCGGCGCGGACGCGAGCGC	63
QY	21	ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	64	CGCTACACCACTCGTCCGCGGACAGGAGGAGGCAAGCCCGCAGAAATCGTACAGC	123
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	124	TCCAGCGAGACCCCTGAAGGCCCTACGACCAAGACGCCCGCTAGCCTATGGCAGCCGCGTC	183
QY	61	LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu	80
Db	184	AAGGACATTGTGCCGCGAGGAGGCCGAGGAATTCTGCCGACACAGGTGCCAACTTCACCCCTG	243
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	244	CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCGTACCGGACAGACATT	303
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	304	GGCCTCCCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	363
QY	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140
Db	364	ACGGTGTGTCCCCTGAGCACCCCGTGGTCTGTGTGGGGCGGAGCACACGGTCCAGGGCGC	423
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	424	AGTCTCTGCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA	483
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	484	AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCAACGCGGGCTCCGGACGCCCGCG	543
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200
Db	544	CCGCGGCTCTGCACGCCCAACACCCCAACCCAGCACCAACGCGGGCTCCATTAACTCCCTG	603
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	604	AACCGGGCAACTTCACGCCGAGGAGCAACCCCAACCCAGCCCGCCCAACCGGACCACTCGCTC	663
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTyrLeu	240
Db	664	TCGGAGAGCCCCCTGCCGGCGCGCCAGGAGCCTGCCACGCGCCAGGAGAACTGGCTG	723
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
Db	724	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA	783
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	784	TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGCTTAC	843
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	844	AGTGACGGGCACTTCTCTTCAAGCCTTGAGGACCTCCCGGCTCTTCTGCACCAATCA	903
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320
Db	904	CCAGGGTACCCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCTGCCCGCG	963
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTyrLys	340

Db	964	AGCACCTTCGCCTGGCGGCTTTAACTCAAGAGAGCCCTCCAAGTACTGTAACTGGAAG	1023
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe	360
Db	1024	TGCGCAGCCCTGAGCGCATCGTCATCTCAGCCACTCTGGTCTATCTGCTGGCATACTT	1083
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet---	379
Db	1084	GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGACG	1143
QY	380	----TyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe	398
Db	1144	GATTATGAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCCCT	1203
QY	398	uTyrPro 400	
Db	1204	ATACCCC 1210	
RESULT 11			
US-10-290-578-11			
; Sequence 11, Application US/10290578			
; Publication No. US20030078389A1			
; GENERAL INFORMATION:			
; APPLICANT: Schaefer, Gabriele M.			
; Sliwkowski, Mark			
; TITLE OF INVENTION: Gamma-Heregulin			
; NUMBER OF SEQUENCES: 11			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Genentech, Inc.			
; STREET: 460 Point San Bruno Blvd			
; CITY: South San Francisco			
; STATE: California			
; COUNTRY: USA			
; ZIP: 94080			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: WinPatin (Genentech)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/290,578			
; FILING DATE: 08-No. US20030078389A1-2002			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/08/891,845			
; FILING DATE: <Unknown>			
; APPLICATION NUMBER: 60/021640			
; FILING DATE: 12-Jul-96			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Lee, Wendy M.			
; REGISTRATION NUMBER: 40,378			
; REFERENCE/DOCKET NUMBER: P1043			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415/225-1994			
; TELEFAX: 415/952-9881			
; TELEX: 910/371-7168			
; INFORMATION FOR SEQ ID NO: 11:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2387 base pairs			
; TYPE: Nucleic Acid			
; STRANDEDNESS: Single			
; TOPOLOGY: Linear			
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:			
US-10-290-578-11			
Alignment Scores:			
Pred. No.:	1,77e-120	Length:	2387
Score:	1285.00	Matches:	235
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	59.96%	Indels:	0
DB:	15	Gaps:	0


```
US-10-029-020-14_COPY_1_400 (1-400) x US-10-290-578-11 (1-2387)
QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCCTGCAGAACACACGCGCGCTCCGGACGCCCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACCCCCAACACGACACACGCGCGCTCCATTAATCCCTGAACCGGGCAACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCACGCCGCGCCCCACGACCATCGCTCTCCGAGAGCCCCCT 180
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCCGGCGGCCCGAGGAGCCTGCCACGCCCGAGGAACTGGCTGCTCAACAGCAACATC 240
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGACATTGCGAGCAACCTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGGTTACAGTGACGGGCACTTC 360
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGTGCACCATCACAGGGTACCCACTG 420
QY 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCCAGACAGTGTACTCTCTCGCCCCGACCCCTGCCCGCAGCACCTTCGCCCGG 480
QY 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345
Db 481 CCGGCCTTTAACCTCAAGAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540
QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
Db 541 GCCATCGTCATCTCAGCCACTCTGGTCATCTCTGTGGCATACTTTGTGGCCATGCACCTG 600
QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTTGGCTAAACTGGCACCTGCAGCCGATGGAGGGGAGATGTATGAGATCACGGAGGAC 660
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATATACCC 705

RESULT 12
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication NO. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca

; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.: 2.04e-89 Length: 8645
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best Local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 17 Gaps: 12

US-10-029-020-14_COPY_1_400 (1-400) x US-10-038-854-37 (1-8645)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg--ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAAGAACGCGAGCGCTTACTGCTCCCTGACCCAGACGACGAGAGAAGGAA 210
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACACAAATTCTCTCCGACACAATGAGGAGTGCCTGGGTACCCACACAGAAGTCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCACGCGAGACATTGAAAGCTTTTGATCATGATGATTCCTCGCGGCTGCTTACGGC 330
QY 58 SerArgValLysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGGATTTGGTTTCACAGAGAAGCAGACGAGTTCACTAGACAAGGACAGAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTTACCCTAAGGCAGTTAGGAGTTTGTGAACCAACGACCACTCGAAGAGAGACTGGCATTGTT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
```

Db 451 GCGGAATGGGGCTCCCTCACAGAGGTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTGATGTCCCCAGAGCATGCCATGAGACTTTGGGCGAGGGGTCAAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGGCCGAGCTCCTGCCTGTCAAGTCGGTCCAACTCAGCCCTCACCCCTGACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAGTCCGACAGTGAGAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690
QY 175 ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGCCGCCT-----TCCCATAGCAGCAGCTCTGCACAGCATCAT--- 741
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCTCAACAGAACTCCCTGACCAATAGAGAGAACCCAGAGTCCGGCC 801
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CGGCTGCTTTGCCGCCGAGCTGCAAAACCACA 837
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCCGAGTCCGTCAGCTGCAGGACAGCTGGTCTCTGGCAGTAATGTACCCTGGAAGC 897
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTCCTATTCAAAACA 918
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAACGCCACTGTTCAAGTACTGCAACCCAGGATACACAATGGCATCTGSC 978
QY 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTCAACCACTACTCGCCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTT 1038
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAGAAGTCTTCAAGTACTGTAGCTGGAAATGCACTGCACTGTGTGCCGTACGG 1098
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCCTGGCAATACTCTGTCTTATTTATAGCAATGCATCTCTTTTGGCCTC 1158
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAAATGACACATTTGAGAATGAAAAGTGAATTTCTGAT 1218
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAAACACTGTGTCTATTACCT 1248

RESULT 13

US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J

Alignment Scores:
Pred. No.: 2.05e-89 Length: 8675
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best Local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 17 Gaps: 12

US-10-029-020-14_COPY_1_400 (1-400) x US-10-038-854-35 (1-8675)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAAGAACGACGAGCCTTACTGCTCCCTGACCAAGACGACGAGAGAGGAA 210
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACACAAATTCTCCGACAGACAATGAGGAGTGCCGGGTACCCACACAGAAATCC 270

; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

```
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
|||
Db 271 TACAGTTCACGCGAGACATTGAAAGCTTTTGATCATGATTCCTCGGGCTGCTTACGGC 330

QY 58 SerArgVallLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn 77
:::|||||
Db 331 AACAGAGTGAAGGATTTGGTTACAGAGAAGCAGACGAGTTCACAGACAAGGACAGAAT 390

QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
|||||
Db 391 TTTACCTAAGGCAGTTAGGAGTTTGTGAACCAAGCAACTCGAAGAGGACTGGCATTGTGT 450

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
:::|||||
Db 451 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTCAGGGTCAGATGCTGATACT 510

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
|||
Db 511 GAAAATGAAGCAGTGATGTCCTCCAGAGCATGCCATGAGACTTTGGGGCAGGGGGTCAAA 570

QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
|||||
Db 571 TCAGGCCGAGCTCCTGCCTGCTCAAGTCGGTCCAACCTCAGCCCTCACCCCTGACAGATACG 630

QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
|||||
Db 631 GAGCAGCAAAACAAGTCCGACAGTGAGAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690

QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
|||
Db 691 ACCCTGCAGCCCTTGCCGCCT-----TCCCATAAAGCAGCACTCTGCACAGCATCAT-- 741

QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
|||||
Db 742 CCATCCATCACTTCTCTCAACAGAAACTCCCTGACC AATGAGAAGAACCCAGAGTCCGSCC 801

QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
|||
Db 802 CCG-----CGGCTGCTTTGCGCGCGAGCTGC AAACCACA 837

QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
|||
Db 838 CCGAGTCCGTCAGCTGCAGGACAGCTGGGTCTTGGCAGTAATGTACC ACTGGAAAGC 897

QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
|||
Db 898 AGG----- 900

QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
|||
Db 901 -----CATTTCTTATTC AAAACA 918

QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
|||
Db 919 GGAACAGGTACAACGCCACTGTTTCAGTACTGCAACCC CAGGATACACAATGGCATCTGGC 978

QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
:::|||||
Db 979 TCTGTTATTCACCACCTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTTT 1038

QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
|||
Db 1039 AAATTCAAGAAGTCTTCAAAGTACTGTAGCTGGAATGC ACTGCACTGTGTGCCGTAGGG 1098

QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
:::|||||
Db 1099 GTCTCGGTGCTCCTGGCAATACTCCTGTCTATTATTTATAGCAATGCATCTCTTTGGCCTC 1158

QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
|||
Db 1159 AACTGGCAGCTACAGCAGACTGAA AATGACACATTTGAGAATGGBAAAAGTGAATTCTGAT 1218

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
```

```
Db 1219 ACC-----ATGCCAAACAACACTGTGTCA TTACCT 1248

RESULT 14
US-10-144-194A-81
; Sequence 81, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 9695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (435)..(8336)
US-10-144-194A-81

Alignment Scores:
Pred. No.: 2.2e-85 Length: 9695
Score: 948.00 Matches: 207
Percent Similarity: 64.06% Conservative: 55
Best Local Similarity: 50.61% Mismatches: 91
Query Match: 44.24% Indels: 56
DB: 16 Gaps: 13

US-10-029-020-14_COPY_1_400 (1-400) x US-10-144-194A-81 (1-9695)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
|||
Db 435 ATGGATGTAAAGGACCGCGCA--CACCGCTCTTTGACCAGGACGCTGTGGCAAAGAG 491

QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
|||||
Db 492 TGTCGCTACACAAGCTCCTCTCTGGACAGTGAGGACTGCCGCGTGCCCCACACAGAAATCC 551

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
|||||
Db 552 TACAGTCCAGTGAGACTCTGAAGGCCTATGACCATGACAGAGGATGC ACTATGGAAAC 611

QY 59 ArgVallLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
|||||
Db 612 CGAGTCACAGACCTCATCCACCGGAGTCAGATGAGTTTCTTAGACAAGGAACCAACTTC 671

QY 79 ThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThr 98
|||||
Db 672 ACCCTTGCCGAACCTGCGCATCTGTGAGCCCTCC---CCACACCGAAGCGGCTACTGTCTCC 728

QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
|||||
Db 729 GACATGGGATCCTTCACCAGGGCTACTCCCTTAGCACAGGCTCTGACGCCGACTCCGAC 788

QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
:::
Db 789 ACCGAGGAGGGATGTCTCCAGAACACGCCCATCAGACTGTGGGCAGAGGGATATAAATCC 848

QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
|||||
Db 849 AGGCGAGTTCGGCCTGTCCAGTCGTGAAAACCTCGGCCCTTACCCTGACTGACTCTGAC 908

QY 159 HisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArg-----Leu 176
:::|||||
Db 909 AACGAAAAACAAATCAGATGATGAGAAACGGCCCTCCGAAACCAACACAGCCAGTCGACTCTG 968

QY 177 ArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAsnSer 196
|||
Db 969 AGGCCCCCTCTCCCAACCC---CCTCACAAACCAACACGCTGTCCCATCACCCAC---TCGTCC 1022
```



```
QY 197 IleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn-----ProSer 212
Db 1023 GCCAACTCCCTCAACAGGAACCTCACTGACCAATCGGCGAGTCAGATCCACGCCCGGCC 1082
QY 213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro 232
Db 1083 CCAGCGCCCAATGAC-----CTGCCACCACACCA-----GAGTCC 1118
QY 233 AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu 252
Db 1119 GTTCAGCTTCAGACAGCTGGGTGCTAAACAGCAACGTCGCCACTGGAGACCCGG----- 1172
QY 253 GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly 272
Db 1172 ----- 1172
QY 273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys---ProGlyGly 291
Db 1173 -----CACTTCCTCTCAAGACCTCCTCGGGG 1199
QY 292 ThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyr 311
Db 1200 AGCACACCTTGTTTCAGCAGCTCTTCCCGGGATACCCCTTTGACTCAGGAACGGTTTAC 1259
QY 312 SerProProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLys 331
Db 1260 ACGCCCCCGCCCCCTGCTGCCCGAGGAATCTTCTCCAGGAAGGCTTCAAGCTGAAG 1319
QY 332 LysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAla 351
Db 1320 AAGCCCTCCAAATACTGCAGCTGGAATGTGTGCCCTCTCCGCCATTGCCCGGCCCTC 1379
QY 352 ThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHis 371
Db 1380 CTCTTGGCTATTTTGTGGCGTATTTCATAGCAATGCATCTGCTCGGACTCAATTGGCAA 1439
QY 372 LeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrpPro 391
Db 1440 CTCCAGCCTGCAGATGGGCACACCTTT-----AACAAATGGGATAAGGACCGGCTTACCA 1493
QY 392 ValProThrAspValSerLeuTyrPro 400
Db 1494 GGAACGATGATGTGGCAACAATGCCA 1520

RESULT 15
US-09-808-602-78
; Sequence 78, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-808-602-78
```

```
Alignment Scores:
Pred. No.: 1.34e-82 Length: 8689
Score: 920.00 Matches: 211
Percent Similarity: 54.91% Conservative: 52
Best Local Similarity: 44.05% Mismatches: 90
Query Match: 42.93% Indels: 126
DB: 9 Gaps: 14

US-10-029-020-14_COPY_1_400 (1-400) x US-09-808-602-78 (1-8689)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArg---AspAlaGlu 19
Db 133 ATGGATGTGAAGGATCGGCGA---CATCGCTCTTTGACCAGGGGACGGTGTGGCAAGGAG 189
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 190 TGTGCTACACACAGCTCCTCTCTGGACAGTGAGGACTGCCGTGTGCCACGCAAGTCC 249
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 250 TACAGTTCAGTGAGACCCCTGAAGGCTTATGACCATGACAGCAGAATGCACTATGGAAAC 309
QY 59 ArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhe 78
Db 310 CGAGTCAAGACCTGGTGCCACCGGAGTCCGATGAGTTTTCTAGACAAGGGGCTAATTTC 369
QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThr 98
Db 370 ACCCTGGCAGAATTGGGAATCTGCGAGCCCTCC---CCACACCGAAGTGGTTACTGTTC 426
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 427 GACATGGGGATCCTCCACCAGGCTACTCCTGAGCACTGGGTCTGATCGGACTCGGAC 486
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
Db 487 ACCGAGGGAGGATGTCTCCAGAACATGCCATCAGACTGTGGGACGAGGATAAAATCG 546
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 547 AGCGCAGCTCTGGCTTGTCCAGCCGCGAGAACTCAGCCCTTACTCTGACTGATTCTGAC 606
QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
Db 607 AATGAAAAATAAATCGGATGACGACAATGGTCGACCCATTCACCTACATCCTCGTCTAGC 666
QY 168 ----- 168
Db 667 CTCCTCCCATCTGCTCAGCTGCCTAGCTCCCATATAATCCTCCACAGTTAGCTGCCAGATG 726
QY 168 ----- 168
Db 727 CCATTGCTAGACAGCAACACCTCCCATCAGATCATGAGACACCAACCCCGATGAGGAATTC 786
QY 169 -----Gly 169
Db 787 TCCCTTAATTTCATACCTGCTCAGAGCATGCTCAGGGCCCCAGCAAGCCTCCAGTAGTGGC 846
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla 186
Db 847 CCTCCGAACCCACACAGCCAGTCAACGCTGAGGGCCCCCTCTGCCACCT---CCTCATAAC 903
QY 187 HisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
Db 904 CACACCTGTCCCACCACCAC---TCCTCTGCCAACTCCCTCAACAGAAACTCACTGACC 960
QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
Db 961 AATCGGGGAGTCAAAATCCACGCCCGCCAGCTCCTGACCCCAATGAC-----CTGGCCACC 1014
QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsn 242
```

Db 1015 ACGCCG-----GAGTCCGTTACAGCTCCAGGACAGCTGGGTGCTGAAC 1056

Qy 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThrLeuGln 262
|||||:::|||||

Db 1057 AGTAACGTGCCGCTGGAGACGCGG----- 1080

Qy 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282

Db 1080 ----- 1080

Qy 283 GlyHisPheLeuPheLys--ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
|||||:::|||||

Db 1081 ---CACTTCCTCTTCAAGACGCTCCTCCGGAAGCACACCCCTGTTTCAGCAGCTCTTCTCCA 1137

Qy 302 GlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArgSer 321
|||||:::|||||

Db 1138 GGATACCCCTTGACCTCAGGGACCGTTTATACACCACCCCGCTGCTGCCACGGAAT 1197

Qy 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys 341
|||||:::|||||

Db 1198 ACATTCTTAGGAAGGCCTTCAAGCTGAAGAAACCCCTCCAAATACTGCAGITGGAAATGC 1257

Qy 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361
|||||:::|||||

Db 1258 GCCGCCCTGTCTGCCATTGCCGCTGCCCTCCTTCTGGCCATTTTGTGGCCTATTTCATA 1317

Qy 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu 381
|||||:::|||||

Db 1318 GCAATGCATCTGCTCGGACTCAATTGGCAACTCCAGCCGCGCAGATGGACACACCTTT--- 1374

Qy 382 IleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
::: ||| |||||::: |||

Db 1375 ---AACAAATGGCGTAAGGACCGGCTTACCAGGAACGATGATGTGGCAACAGTGCCA 1428

This Page Blank (usp

This Page Blank (uspic

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 118.339 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400

Perfect score: 2143

Sequence: 1 MDVKRKPYSLTRRDAER.....EITEDTASSWPVPTDVSLYP 400

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_258@runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2138	99.8	1680	3	US-08-891-845-3
2	2138	99.8	1680	4	US-09-514-573-3
3	2138	99.8	3111	3	US-08-891-845-1
4	2138	99.8	3111	4	US-09-514-573-1
5	1285	60.0	2387	3	US-08-891-845-11
6	1285	60.0	2387	4	US-09-514-573-11
7	138	6.4	4446	4	US-09-231-899-69
C 8	131.5	6.1	1860	4	US-09-051-994-1
9	131.5	6.1	2949	4	US-09-623-326-6
C 10	130.5	6.1	4403765	3	US-09-103-840A-2
C 11	130.5	6.1	4411529	3	US-09-103-840A-1
12	130	6.1	1998	4	US-09-252-991A-11748
					Sequence 11748, A

C 13	130	6.1	3180	4	US-09-252-991A-11618	Sequence 11618, A
C 14	128.5	6.0	8438	1	US-07-945-283-1	Sequence 1, Appli
15	127.5	5.9	4456	4	US-09-095-443-1	Sequence 1, Appli
C 16	127	5.9	71989	4	US-09-443-501A-2	Sequence 2, Appli
C 17	126.5	5.9	43280	2	US-08-804-227C-1	Sequence 1, Appli
18	125	5.8	2507	2	US-08-471-066B-1	Sequence 1, Appli
19	124.5	5.8	2733	4	US-09-623-326-2	Sequence 2, Appli
20	124.5	5.8	5121	4	US-09-252-991A-15189	Sequence 15189, A
21	122.5	5.7	2499	4	US-09-758-282B-100	Sequence 100, App
22	122.5	5.7	2517	4	US-09-758-282B-147	Sequence 147, App
23	122.5	5.7	2517	4	US-09-758-282B-149	Sequence 149, App
24	122.5	5.7	2517	4	US-09-758-282B-152	Sequence 152, App
25	122.5	5.7	2517	4	US-09-758-282B-168	Sequence 168, App
26	122.5	5.7	2517	4	US-09-758-282B-171	Sequence 171, App
27	122.5	5.7	2517	4	US-09-758-282B-180	Sequence 180, App
28	122.5	5.7	2517	4	US-09-758-282B-183	Sequence 183, App
29	122.5	5.7	2517	4	US-09-758-282B-186	Sequence 186, App
30	122.5	5.7	2733	4	US-09-623-326-1	Sequence 1, Appli
31	122.5	5.7	4287	1	US-08-244-189-1	Sequence 1, Appli
32	122.5	5.7	4287	1	US-08-306-691B-53	Sequence 53, Appli
C 33	122.5	5.7	50937	3	US-09-428-517-1	Sequence 1, Appli
34	122	5.7	4689	3	US-09-105-537-34	Sequence 34, Appli
35	122	5.7	36778	3	US-09-105-537-5	Sequence 5, Appli
36	122	5.7	38506	3	US-09-320-878-19	Sequence 19, Appli
37	122	5.7	38506	4	US-09-141-908-1	Sequence 1, Appli
38	122	5.7	38506	4	US-09-657-440-19	Sequence 19, Appli
39	121.5	5.7	1686	2	US-08-648-657-1	Sequence 1, Appli
40	121.5	5.7	1689	2	US-08-648-657-2	Sequence 2, Appli
41	121.5	5.7	2517	4	US-09-758-282B-174	Sequence 174, App
42	121.5	5.7	2727	4	US-09-623-326-4	Sequence 4, Appli
43	121.5	5.7	2943	4	US-09-379-530B-3	Sequence 3, Appli
44	121	5.6	4078	4	US-09-016-434-1152	Sequence 1152, Ap
45	121	5.6	4553	4	US-09-023-655-975	Sequence 975, App

ALIGNMENTS

RESULT 1

US-08-891-845-3
; Sequence 3, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPattn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-891-845-3

Alignment Scores:

Pred. No.: 2.49e-187 Length: 1680
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-08-891-845-3 (1-1680)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 1 ATGGACGTGAAGGAGAGGAAAGCCTTACCGCTCGTACCCGCGCGGCGGACGCGGAGCGC 60
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 61 CGGTACACAGCTCGTCCGCGGACAGCGAGGAGGCAAGCCCGCAGAAATCGTACAGC 120
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 121 TCCAGCGAGACCCCTGAAGCCTACGACGAGACGCCCGCTAGCTATGGCAGCGCGTC 180
QY 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 181 AAGGACATGTGCCGCGCAGGAGGCGGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG 240
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 241 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCGCTCACGGGACCCCTGTACCGGACAGACATT 300
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 301 GGCTCCCCCTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 360
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 361 ACGGTGCTGTCCCTGAGCACCCCGTGCCTGTGGGGCGGAGCACACGGTCAAGGCGC 420
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 421 AGCTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACATCACCGACACCGAGCATGAA 480
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 481 AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCAACGGCGGCTCCGGACGCCGCG 540
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 541 CCGCCGCTCTCGCACGCCCCACACCCCAACAGCACCAACCGCGGCTCCATTAACTCCCTG 600
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 601 AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGCGGCGGCGGCGGACCACTCGCTC 660
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 661 TCCGGAGAGCCCGCTGCGGGCGGCGGCGGAGAGCTGCCACGCCCGCAGGAACTGGCTG 720
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 721 CTCAACAGCAACATCCCCCTGGAGACCAGAACTAGGCAAGCAGCCATTCCTAGGGACA 780
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 781 TTGAGGACACCTCATTTGATGGACATTTCTGGCGGCTCCCGCCATGATGGGCTTAC 840

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 841 AGTGACGGGCACCTTCCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCATCA 900
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 901 CCAGGTACCCACTGACGTCAGACACAGTGACTCTCTCTCGGCCCCGACCCCTGCCCGC 960
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 961 AGCACCTTCGCCCGCGGCGCTTTAACCTCAAGAAAGCCCTCCAAAGTACTGTAACTGGAAG 1020
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1021 TGGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTT 1080
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1081 GTGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1140
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1141 GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCCCTATACCCC 1200

RESULT 2

US-09-514-573-3
; Sequence 3, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-3

Alignment Scores: 2.49e-187 Length: 1680
Pred. No.: 2.49e-187

Score:	2138.00	Matches:	399
Percent Similarity:	99.75%	Conservative:	0
Best Local Similarity:	99.75%	Mismatches:	1
Query Match:	99.77%	Indels:	0
DB:	4	Gaps:	0
US-10-029-020-14_COPY_1_400 (1-400) x US-09-514-573-3 (1-1680)			
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	1	ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCGACGCCGAGCGC	60
QY	21	ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	61	CGCTACACCAGCTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC	120
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	121	TCCAGCGAGACCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC	180
QY	61	LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu	80
Db	181	AAGGACATTGTGCCGCGAGGAGGCGGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG	240
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	241	CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCGTGTACCGGACAGACATT	300
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	301	GGCTCCCCCACTGCGGCTACTTCCATGGGGGCTGGTCTGTATGCCGACATGGAGGCTGAC	360
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	361	ACGTGTCTGCCCTGAGCACCCCGTGCCTGTGTGGGCGCGAGCACACGGTCAGGGCGC	420
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	421	AGTCTCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA	480
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	481	AACACTGAGACTGATCATCCGGGCGGCTGCAGAACCCACGCGGGCTCCGGACGCCGCG	540
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200
Db	541	CCGCGCTCTCGACGCCACACCCCAACACGACCCAGCCAGCCGCGCTCCATTAACTCCCTG	600
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	601	AACCGGGGCAACTTCACGCCGAGGAGCAACCCCAAGCCCGCGGCCACGGACCACTCGCTC	660
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	661	TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCCCAGGAGAACTGGCTG	720
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
Db	721	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA	780
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	781	TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGGGCCCTCCGCCCATGATGGGGCTTAC	840
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	841	AGTGACGGGCACTTCTCTTCAAGCCTGGAGGCACCTCCCGGCTCTTCTGCACCACATCA	900
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320
Db	901	CCAGGGTACCCCACTGACGTCCAGCACAGTGACTCTCTCTCCGCCCCGACCCCTGCCCCGC	960
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340

Db	961	AGCACCTTCGCCCGCGCCTTTAAACCTCAAGAAGCCCTCAAGTACTGTAACCTGAAG	1020
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1021	TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCCTGCTGGCATACTTT	1080
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1081	GTGGCCATGCACCTGTTTGGCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT	1140
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	1141	GAGATCAGGAGGACACAGCCAGCTTGGCCTGTGCCAACCGACGTCTCCCTATATACCC	1200
RESULT 3			
US-08-891-845-1			
; Sequence 1, Application US/08891845			
; Patent No. 6096873			
; GENERAL INFORMATION:			
; APPLICANT: Schaefer, Gabriele M.			
; APPLICANT: Sliwowski, Mark			
; TITLE OF INVENTION: Gamma-Heregulin			
; NUMBER OF SEQUENCES: 11			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Genentech, Inc.			
; STREET: 460 Point San Bruno Blvd			
; CITY: South San Francisco			
; STATE: California			
; COUNTRY: USA			
; ZIP: 94080			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: WinPatin (Genentech)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/891,845			
; FILING DATE:			
; CLASSIFICATION: 435			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 60/021640			
; FILING DATE: 07/12/96			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Lee, Wendy M.			
; REGISTRATION NUMBER: 40,378			
; REFERENCE/DOCKET NUMBER: P1043			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415/225-1994			
; TELEFAX: 415/952-9881			
; TELEX: 910/371-7168			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 3111 base pairs			
; TYPE: Nucleic Acid			
; STRANDEDNESS: Single			
; TOPOLOGY: Linear			
US-08-891-845-1			
Alignment Scores:			
Pred. No.:	6.58e-187	Length:	3111
Score:	2138.00	Matches:	399
Percent Similarity:	99.75%	Conservative:	0
Best Local Similarity:	99.75%	Mismatches:	1
Query Match:	99.77%	Indels:	0
DB:	3	Gaps:	0
US-10-029-020-14_COPY_1_400 (1-400) x US-08-891-845-1 (1-3111)			
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	334	ATGACGTGAAGGAGAGAACCTTACCGCTCGCTGACCCGGCGCCGACGCCGAGCGC	393

Qy 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCAAGTCTCGTCGGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC 453

Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGCCCTACGACCAAGCGCCCGCTAGCCTATGGCAGCCGCGTC 513

Qy 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 514 AAGGACATTGTGCCGACGAGGCGGAGGAATTCTGCCGACAGGTGCCAATCTCACCCCTG 573

Qy 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAGAAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 633

Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCCACTGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGCTGAC 693

Qy 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGTGTCTGTCCCCTGAGCACCCCGTGCCTGTGTGGGCGGAGCACACGGTCAAGGCGC 753

Qy 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCCTGCCTGTCCAGCGGGCCCAATTCCAATCTCACACTCACGACACCGAGCATGAA 813

Qy 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGGCGCTGCAGAACCAACCGCGGCTCCGACGCGCGC 873

Qy 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTGCACGCCACACCCCAACCAACGACCAACCGCGGCTCCATTAACCTCCCTG 933

Qy 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGGCCCCACGGACCACTCGCTC 993

Qy 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCCCTGCCGGCGGCCAGGAGCTGCCACGCCAGGAGAACTGGCTG 1053

Qy 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAAGAACCTAGGCAAGCAGCCATTCTTAGGGACA 1113

Qy 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCTCCCGCCATGATGGGCTTAC 1173

Qy 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPhePheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCCCGTCTTCTGCACCACATCA 1233

Qy 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTACCCCACTGACGCTCAGACAGTGACTCTCTCCGCCCCGACCCCTGCCCCGC 1293

Qy 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCCCGCCCTTTAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1353

Qy 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGTGCATACCTT 1413

Qy 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCATGCACCTGTTTGGCCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1473

Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCAGCGTCTCCCTATACCCC 1533

RESULT 4

US-09-514-573-1

; Sequence 1, Application US/09514573

; Patent No. 6500941

; GENERAL INFORMATION:

; APPLICANT: Schaefer, Gabriele M.

; APPLICANT: Sliwowski, Mark

; TITLE OF INVENTION: Gamma-Herregulin

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/514,573

; FILING DATE: 28 FEB 2000

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/891845

; FILING DATE: 10 JULY 1997

; APPLICATION NUMBER: 60/021640

; FILING DATE: 07/12/96

; ATTORNEY/AGENT INFORMATION:

; NAME: Lee, Wendy M.

; REGISTRATION NUMBER: 40,378

; REFERENCE/DOCKET NUMBER: P1043

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415/225-1994

; TELEFAX: 415/952-9881

; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3111 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

US-09-514-573-1

Alignment Scores:

Pred. No.: 6.58e-187 Length: 3111

Score: 2138.00 Matches: 399

Percent Similarity: 99.75% Conservative: 0

Best Local Similarity: 99.75% Mismatches: 1

Query Match: 99.77% Indels: 0

DB: 4 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-09-514-573-1 (1-3111)

Qy 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20

Db 334 ATGCACGTGAAGGAGAGGAAAGCTTACCGCTCGCTGACCCCGCGCGCAGCGAGCGC 393

Qy 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40

Db 394 CGTACACCAAGCTCGTCCCGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC 453

Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60

Db 454 TCCAGCGAGACCCCTGAAGGCCCTACGACCAAGCGCCCGCTAGCCTATGGCAGCCGCGTC 513

Qy 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80

Db 514 AAGGACATTGTGCCCGCAGGAGCGCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAGAGTAACGCCCCCTCACGGGACCCTGTACGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCTTCCCCCACTGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGTGTCTCTCCCTGAGCACCCCGTGCCTGTGTGGGCCGGAGCACACGGTCAAGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCTGCAGAACCAACCGCGGCTCCGGACCGCCCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTCGACGCCACACCCCAACAGCACCAACCGCGGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCACCGGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCCCTGCCGGCGGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCATTCCTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCGGCTCTTCTGCACCACATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTACCCCACTGACGTCCAGCACAGTGTAATCTCTCCCGCCCGACCCCTGCCCGCG 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGCGCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGCGCAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCACTCCTGTGTCATCTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTTGGCCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 5

US-08-891-845-11
; Sequence 11, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark

; TITLE OF INVENTION: Gamma-Herectulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-891-845-11

Alignment Scores:

Pred. No.: 1.66e-108 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.96% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x US-08-891-845-11 (1-2387)

QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCTGCAGAACCCAGCGGGCTCCGGACGCCCGCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACCCCCAACCCAGCACCCAGCGGCTCCATTAACTCCCTGAACCGGGCACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCAGCCCGCCCCCACGACCACTCGCTCTCCGAGAGCCCCCT 180
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCCGGCGGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGTGTCTCAACAGCAACATC 240
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCCTGAGACCCAGAAACCTAGCAAGCAGCCATTCTCTAGGACATTGCAGGACCACTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGGACATTCTCGGCGCCTCCCGCATGATGGGGCTTACAGTGACGGGCACTTC 360
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305

Db 361 CTCTTCAAGCCTGGAGGACCTCCCGGCTCTTCTGCACCACATCACAGGGTACCCACTG 420
Qy 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCCAGCACAGTGTACTCTCTCCGCCCGGACCCCTGCCCGCAGCACCTTCGCCCGG 480
Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTyrLysCysAlaAlaLeuSer 345
Db 481 CCGGCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
Db 541 GCCATCGTCATCTCAGCCACTCTGGTCATCTGCTGGCATACTTTGTGGCCATGCACCTG 600
Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTTGGCCTTAACTGGCACCTGCAGCCGATGGAGGGGAGATGTATGAGATCACGGAGGAC 660
Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 705

RESULT 6

US-09-514-573-11
; Sequence 11, Application US/095114573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Slikowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-11

Alignment Scores:
Pred. No.: 1.66e-108 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 59.96% Indels: 0
DB: 4 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x US-09-514-573-11 (1-2387)
Qy 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCGCTGCAGAACCCACGCGGGCTCCGGACGCCCGCGCGCTCTCGCAC 60
Qy 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCCACACCCCAACCAGCACACCACGCGGCTCCATTAACTCCCTGAACCGGGCAACTTC 120
Qy 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCAGCCCGGCCCCCAGGACCACTCGTCTCTCGGAGAGCCCCCT 180
Qy 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuAsnSerAsnIle 245
Db 181 GCCGGCGGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTGCTCAACAGCAACATC 240
Qy 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAsnLeu 265
Db 241 CCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGACATTGCAGGACAACCTC 300
Qy 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGCTTACAGTGACGGGCACTTC 360
Qy 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 361 CTCCTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCACATCACAGGTTACCCACTG 420
Qy 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325
Db 421 ACGTCCAGCACAGTGTACTCTCTCCGCCCGGACCCCTGCCCGCAGCACCTTCGCCCGG 480
Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345
Db 481 CCGGCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540
Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
Db 541 GCCATCGTCATCTCAGCCACTCTGGTCATCTCTGTGGCATACTTTGTGGCCATGCACCTG 600
Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
Db 601 TTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGAGATGTATGAGATCACGGAGGAC 660
Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 661 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 705

RESULT 7

US-09-231-899-69
; Sequence 69, Application US/092311899
; Patent No. 6566583
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Metz, James G
; APPLICANT: Facciotti, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES
; FILE REFERENCE: CGNE.131.02US
; CURRENT APPLICATION NUMBER: US/09/231,899
; CURRENT FILING DATE: 1999-01-14
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; EARLIER APPLICATION NUMBER: 09/090,793
; EARLIER FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 69
; LENGTH: 4446

; TYPE: DNA
; ORGANISM: Schizochytrium aggregatum
US-09-231-899-69
Alignment Scores:
Pred. No.: 0.0208 Length: 4446
Score: 138.00 Matches: 106
Percent Similarity: 34.18% Conservative: 42
Best Local Similarity: 24.48% Mismatches: 163
Query Match: 6.44% Indels: 122
DB: 4 Gaps: 20
US-10-029-020-14_COPY_1_400 (1-400) x US-09-231-899-69 (1-4446)
QY 19 GluArgArgTyrThrSerSerSerAlaaspSerGluGluGlyLysAlaProGlnLysSer 38
Db 3040 GAGGTGCGCAGCTCTATTGCTGCGATTGAAGCGCTCGGCGGCAAGGCC-----ATC 3090
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAsp-----GlnAspAlaArg 53
Db 3091 TACTCGTGTGCGACGTGAACCTCTGCGCCGACGTGGCCAAAGGCCGTGCGCGATGCCGAG 3150
QY 54 LeuAlaTyrGlySerArgValLysAspIleValProGlnGluAla----- 68
Db 3151 TCCAGCTCGGTGCCCGCTCTCGGCGATCGTTTCATGCTCGGGCGTGTCTCGGACCGT 3210
QY 69 -----GluGluPheCysArgThrGlyAlaAsnPheThrLeuArg 81
Db 3211 CTCATCGAGAAGAAGCTCCCCGACGAGTTC-----GACGCCGTCTTTGGCACCAAG 3261
QY 82 GluLeuGlyLeuGluGluVal-ThrProProHisGlyThrLeuTyrArgThrAspIleGl 101
Db 3262 GTCAACCGTCTCGAGAACCTCCTCGCCGCG-----TCGACCGC 3300
QY 101 yLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaaspMetGluAlaAspTh 121
Db 3301 GCCAACCTCAAGCACATGGTCTCTTCAGTCTCGTCTCGCCGCTTCACGGCAACGTGCGC 3360
QY 121 rValLeuSer-----ProGluHisProValArgLeuTrpGly-----Ar 134
Db 3361 CAGTCTGACTAGCCATGGCCCAACGAGGCCCTTAACAAGATGGGCTCGAGCTCGCCAAAG 3420
QY 134 gSerThrArgSerGlyArgSerSerCysLeuSer-SerArgAlaAsnSerAsnLeuThrL 154
Db 3421 GACGTCTCGTCAAGTCGATCTGCTTCGTCCTCGTGACGGTGGCATGGTGACGCCGCGAG 3480
QY 154 euThrAspThrGluHisGluAsnThrGluThrAspHisPro----- 167
Db 3481 CTCAAGAAGCAGTTCAGGAGATGGCGTGCAGATCATCCCCCGAGGGCGGCGTGTAT 3540
QY 168 -----GlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHisA 186
Db 3541 ACCGTGGCGG-----CATCGTGTCTCGGCTCCTCGCCGGTGTAGATCCTTGTGCGGAAC 3594
QY 186 laHisThrProAsnGln-----HisHisAlaAla----- 195
Db 3595 TGGCGCACCCCGTCCAAGAAGGTGCGGTGCGACACCATCACCTGACCCGCAAGATTTC 3654
QY 196 --SerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAlap 215
Db 3655 GCCAAGTCCAACCCCTTCCTCGAGGA-----CCACGT---CATCCAGGGCGCGCGC 3702
QY 215 roThrAspHis----- 218
Db 3703 GTGCTGCCCATGACGCTGGCCATTGGTTCGTCGCGGAGACCTGCCTCGGCCCTCTTCCCC 3762
QY 218 ----- 218
Db 3763 GGCTACTCGCTCTGGGCCATTGACGACGCCCGCAGCTCTTCAAGGGTGTCACTGTGACCGC 3822
QY 219 -----SerLeuSerGlyGluPro---ProAlaGlyGlyAla-----GlnGlup 232
Db 3823 GACGTCAACTGGGAGGTGACCCCTCACCCCGTCGACGGCGCCCTCGGGCGCGGTCAACGTC 3882

QY 232 roAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArg----- 250
Db 3883 CAGGCACAGCTCAAGACCTTTTCCAGCGGCAAGCTGGTCCCGGCTACCGCGCGTCAATC 3942
QY 251 -----AsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGlu 267
Db 3943 GTGCTCTCCAACACGAGGCGGCGCCCGCCCAACGCCACCATGCAAGCGCCCTCGCTCGAT 4002
QY 268 MetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPhe 287
Db 4003 GCCGAT-----CCGGCGCTCCAGGGTCCGTCCTACGACGGCAAGACCTCTTC 4050
QY 288 Lys---ProGly-GlyThrSerProLeuPheCysThrThrThrSerProGlyTyrProLeuTh 306
Db 4051 CACGGCCCGCCTTCCGGGGCATCGATGACGTGCTCTCGTGCACCAAGAGCCAGCTTGTG 4110
QY 306 rSerSerThrValTyrSerProProProArgPro-----LeuProArgSerTh 322
Db 4111 GCCAAGTGCAGCGTGTCCCGGGTCCGACGCCGCTCGCGGCGAGTTTGCCACGGACACT 4170
QY 322 rPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAl 342
Db 4171 GACGCCCATGACCCCTTCGTGAACGACCTGGCCTTTTCAGGCCATGCTCGTCTGGGTGCGC 4230
QY 342 aAlaLeuSerAlaIleValIleSerAlaThr 352
Db 4231 CGCACGCTCGGCCAGGCTGCGCTCCCCAACT 4261

RESULT 8
US-09-051-994-1/c
; Sequence 1, Application US/09051994A
; Patent No. 6602683
; GENERAL INFORMATION:
; APPLICANT: REGENERON PHARMACEUTICALS, INC.
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE EPH FAMILY LIGANDS
; FILE REFERENCE: REG-341-PCT-US
; CURRENT APPLICATION NUMBER: US/09/051,994A
; CURRENT FILING DATE: 1998-04-24
; EARLIER APPLICATION NUMBER: PCT/US96/17201
; EARLIER FILING DATE: 1996-10-25
; EARLIER APPLICATION NUMBER: 60/007,015
; EARLIER FILING DATE: 1995-10-25
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (202)..(1224)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1717)
; OTHER INFORMATION: n=a, c, g, or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1741)
; OTHER INFORMATION: n=a, c, g or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1782)
; OTHER INFORMATION: n=a, c, g, or t
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1795)
; OTHER INFORMATION: n=a, c, g, or t
US-09-051-994-1

Alignment Scores:
Pred. No.: 0.0209 Length: 1860
Score: 131.50 Matches: 113

Percent Similarity: 32.35% Conservative: 41
Best Local Similarity: 23.74% Mismatches: 173
Query Match: 6.14% Indels: 150
DB: 4 Gaps: 24

US-10-029-020-14_COPY_1_400 (1-400) x US-09-051-994-1 (1-1860)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 1648 CTAGAGACGGAAGAGAGGAA-----AGAATAAGAGATAGAGAAGAG 1607

QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 1606 AAA-----ACCAGTGAGAGGCAGAAAGAGCGGTGCCAACCCAAAGGTGGGCTGT 1556

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 1555 TCCCGCCCTAGCTCTGGGTGGGTGAGGGGAC----- 1523

QY 61 LysAsp-----IleValProGlnGluAlaGluGluPhe-----CysArgThr 74
Db 1522 AAGGATGCTGGTGCATGGTCAGAGAGGAGGCTGAGCCTGTGCCCTCCCCAAGGAAAA 1463

QY 75 GlyAlaAsnPheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThr 94
Db 1462 GGATCCAGGGACACTGAGACAGAGGGGGC-----ACCCATGGCCAA 1421

QY 95 LeuTyrArgThrAspIleGly-----Leu 102
Db 1420 ACGGAGGGCAGGAAGTGGGGCAGTGGGAATCCTAAGGAATCCTAAAGTGGAGACGAGG 1361

QY 103 ProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu---AlaAspThr 121
Db 1360 ACAGCAGCCGGAGGAGTGAAGGGCTGGGGGGGCACAGGGCCGAGATGTTAGAGGTG 1301

QY 122 ValLeuSerPro-----GluHisProValArgLeuTrp----- 132
Db 1300 TCCCTCAAACCAAGGAATTAACTGGAGGAGACACCCCAAGAGGCTGGATTCAGGATAGC 1241

QY 133 ---GlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsn 151
Db 1240 CACGTGAGAGAGCCCTCATACCTTGTAGTAGATGTTGGAGGGCTCTGGGGGGGCCCAT 1181

QY 152 ---LeuThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeu 170
Db 1180 CCTGCACGATATACACAGGATGCCCATAGTCACCACTACCTTCTATAGTGGGGCAGA 1121

QY 171 GlnAsnHisAlaArgLeuArgThrPro-----Pro 180
Db 1120 AGGGGGATCTGCAGCCCGCCACCCGACAGAGCTATCCCTAGTCCCGCAGGCTCAGCCT 1061

QY 181 ProProLeu---SerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSer 199
Db 1060 CCCAGGTCCCATCCCACTCCACCCCGCCAGGCCCAGAGACCCCTCCCTCCCGAAGGAGC 1001

QY 200 LeuAsnArgGlyAsnPheThrPro-ArgSerAsnProSerProAlaProThrAspHis 219
Db 1000 CAGGACCAGGTGGCGACTCTCCGAAGGCTTGGCCCGCG----- 961

QY 219 rLeuSerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTr 239
Db 960 -----TCTCCGCCAACACATGGCAGCCCGCCAGCCCTGCCACGCCCAAGAGCA 911

QY 239 pLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuG1 259
Db 910 GCGCCAGCCCCCTGCTGCCCGCCAGCCACTGCAGGCATGCTGGG----- 868

QY 259 yThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAl 279
Db 867 -----AGGGGCAGGGGGC 854

QY 279 aTyrSerAspGlyHisPheLeuPheLysProGlyGly-ThrSerProLeuPheCysThrT 299
|:::|

Db 853 CTTCAGC-----ACCCCGGAGGTTGCA-----TTGCTGGTGG 821

QY 299 hrSerProGly---TyrProLeuThrSerSerThrValTyr----- 311

Db 820 GGTCACTGGCAGGTTCTCCTTCCCAAGCTCCAGGCTGTGGGCTGCCCTCGGTCTCTTT 761

QY 312 -----SerProProProArg-----ProL 318

Db 760 CCATGGGCATTCAGACACAGGTTTTCGGGGGACAGCCCCTCCTCGGGACTTTGTCCCA 701

QY 318 euProArgSerThrPheAla-----ArgProAlaPheAsnLeuLysLysP 333

Db 700 CTYGGAGAGACACCTTCATGCCTCTGGTTAGGCACACACCTCCCTGCAGGCTCTCCAGGC 641

QY 333 roSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValleSerAlaThrL 353

Db 640 CCTCCCGG---GTCCCATC-CGATGTGGCAATGATGTAGTAATCGTGTGCGAGCGGAAC 585

QY 353 euValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuG 373

Db 584 TCGTGGCCCGCAGAGATTAGGGCTATACCTCTGGAACCTTGATGGT----- 541

QY 373 lnProMetGluGlyGlnMetTyrGluIleThrGluAsp----- 385

Db 540 --GAAGCGGAGATCCAGGTCTGGCGGATCACAAGTGAGAAGGAGTTTGGGGCAGGGGT 483

QY 386 -----ThrAlaSerSerTrpProValProThrAspValSerLeu 398

Db 482 GCCTCACAGCGCGCGCCTGAGCACCCCTACCAGGTACAGCTT 439

RESULT 9

US-09-623-326-6

; Sequence 6, Application US/09623326

; Patent No. 6607883

; GENERAL INFORMATION:

; APPLICANT: Frey et al.

; TITLE OF INVENTION: Polymerase Chimerae

; FILE REFERENCE: 4894

; CURRENT APPLICATION NUMBER: US/09/623,326

; CURRENT FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: DE 198 10 879.6

; PRIOR FILING DATE: 1998-03-13

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 2949

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: polynucleotide

US-09-623-326-6

Alignment Scores:

Pred. No.: 0.0432 Length: 2949

Score: 131.50 Matches: 101

Percent Similarity: 37.61% Conservative: 31

Best Local Similarity: 28.77% Mismatches: 115

Query Match: 6.14% Indels: 104

DB: 4 Gaps: 20

US-10-029-020-14_COPY_1_400 (1-400) x US-09-623-326-6 (1-2949)

QY 3 VallysGluArgLysProTyrArg-----SerLeuThrArgArg 16

Db 1665 GTCAAGCACAGGGAACCTTGTAGAGTGGTCTTACTTAGGAAGGCTACGAAAGAACGA 1724

QY 17 AspAlaGluArg-----ArgTyrThrSerSerSerAlaAspSerGluGluGly 32

Db 1725 AGTAGCTCCAAACAAGCCCAAGTGAAGAGAGGATATCAAAGAAGGCTCAGGGAGAGCTACAC 1784

QY 33 -----LysAlaProGlnLys-SerTyrSerSerSerGluThrLeuLysAlaTyrAs 49

Db 1785 AGGTGGATTCTGTCGCGCTGGACGTGGCCTATCTCAGGGCCTTGTCCCTGAGGTGGCCGA 1844

```
Qy 49 pGlnAspAlaArgLeuAlaTyrGlySerArgVallysAspIleValProGlnGluAlaGl 69
Db 1845 GGAGATCGCCGCGCTC-----GAGGCCGA 1868

Qy 69 uGluPheCysArgThrGlyAlaAsnPheThrLeuArgGluLeuGly--LeuGluGluVa 88
Db 1869 GGTCCTCCGCTGGCCGCCACCCCTTCAACCTCAACTCCCGGACCAGCTGGAAGGGT 1928

Qy 88 lThrProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSe 108
Db 1929 C-----CTCTTT---GACGAGCTAGGGCTTCCCGCCATCGGCAAGAC 1967

Qy 108 rMetGlyAlaGly-----SerAspAlaAspMetGluAlaAspThrValLeuSe 124
Db 1968 GGAGAAACCGGCAAGCGCTCCACGAGCGCCGCTCTCTGGAGGC-----CTCCG 2018

Qy 124 rProGluHisProValArgLeu-TrpGlyArgSerThrArgSerGlyArgSerSerCysL 144
Db 2019 CGAGGCCCAACCC-----ATCGTGGAGAAGATCCTGCAGTACCGGGAGCTCACCA--- 2068

Qy 144 euSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGluT 164
Db 2069 --AGCTGAAGACACCT-----ACATTGACCCCTTGCCGACCTCATCCACC----- 2113

Qy 164 hrAspHisProGly-GlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeu 183
Db 2114 -----CCAGGACGG-----CGCCTCCA--- 2132

Qy 184 SerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGly 203
Db 2133 ---CACCGCTTCAACGAGCGGCCACGGCCACGGCAGGCTAAGTAGCTC----- 2180

Qy 204 AsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGlu 223
Db 2181 -----CGATCCCAACCTCCAGAACATCCCGTCCGACCCCGCTTGGGCAGAG 2228

Qy 224 ProProAlaGlyGlyAlaGlnGluProAlaHis-AlaGlnGluAsnTrpLeuLeuAsnSe 243
Db 2229 GATCCGCGGGC-----CTTCATCGCCGAGGAGGGGTGGCTATTGGTGGC 2273

Qy 243 rAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAs 263
Db 2274 C-----CTGGACTA 2282

Qy 263 pAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGl 283
Db 2283 TAGCCAGATAGACTCAGGTGCTGCC-----CACCTCTCCGCGACGAGAACCT 2333

Qy 283 yHisPheLeuPhe-LysProGlyGlyThrSer-----ProLeuPheCysThrT 299
Db 2334 GATCCGGTCTTCCAGGAGGGCGGGACATCCACGAGACCGCCAGCTGGATGTTCCG 2393

Qy 299 hrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuP 319
Db 2394 CGTCCCGGGAGGCGGTGGACCCCTGTATGCGCGGGCCCAAGACCATCAACTTCGG 2453

Qy 319 roArgSerThrPheAlaArgPro 326
Db 2454 GGTCCTCTACGGCATGTCGGCCC 2476
```

RESULT 10

```
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00

; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 5.47e+03 Length: 4403765
Score: 130.50 Matches: 80
Percent Similarity: 36.72% Conservative: 43
Best Local Similarity: 23.88% Mismatches: 120
Query Match: 6.09% Indels: 92
DB: 3 Gaps: 16

US-10-029-020-14_COPY_1_400 (1-400) x US-09-103-840A-2 (1-4403765)

Qy 32 GlyLysAlaProGlnLysSerTyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp 51
Db 1334300 GGTCGGGTCCCAATCTGCATCAGGGTGCTCAAAATCTCTTAC----- 1334256

Qy 52 AlaArgLeuAlaTyrGlySerArgVallysAspIleValProGlnGluAlaGluGluPhe 71
Db 1334255 -----CACGGCCAAGCCAGCGAGAT-----GACTTT 1334229

Qy 72 CysArgThrGlyAlaAsnPheThrLeuArgGluLeuGlyLeuGluValThrProPro 91
Db 1334228 CGCCGGCAAGTGCAACAACCTCGCGTCGGGAAGCCGGATACGACGTGTT-----CCC 1334175

Qy 92 HisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAla 111
Db 1334174 CGTGGG-----AGAACGGGA-----CGATGGTCTGTTCTCCGTGCC 1334136

Qy 112 GlySer-AspAlaAspMetGluAlaAspThrValLeuSerProGluHisProVal-ArgL 131
Db 1334135 ACCAGCGACGGGACTTTTAC-----CTCGTCCAGCCGAATCCCGAGTCGCGGG 1334085

Qy 131 euTrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSer-----SerArgA 148
Db 1334084 CAAAGCGGATGACATCGGCGAACCGCGCAGCGAGTGCTTGCCTACTACCGTTGAGCAGAT 1334025

Qy 148 laAsnSerAsnLeuThrLeu-ThrAspThrGluHisGluAsnThr----- 162
Db 1334024 CGTCGAGGAACATCGCCTTGAACCTCGGGCCGAGCCAGCAAAATGCCGTCGCGCGGTG 1333965

Qy 163 -----GluThrAspHisProGlyGlyLeu 170
Db 1333964 AGAGCAGCCCATACAGGTCGAGGGCAGGGGACGCGACGCGGCCGCGCGGATCAGCA 1333905

Qy 171 GlnAsnHisAlaArgLeuArgThrProProProLeuSerHisAlaHisThrProAsn 190
Db 1333904 AGCTCGACCCAGCCGACGGGGTGGCCGCCACCTGCAGCAGCGGCCACCGCCGAAC 1333845

Qy 191 Gln-HisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAs 210
Db 1333844 CAAGCGCATCAAAACCGCGCTAATCGCTCCGGCCGCGCGT-----CGGCGGACGC 1333791

Qy 210 nProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyAlaGl 230
Db 1333790 CGCCGAGGACACCGCGCGGCGACCCCGGTC---GGGACGCCC-----GGCGGCGCACG 1333740

Qy 230 nGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrAr 250
Db 1333739 CCAGGGTGTATGGGCCCCCGCCGACAGGCCCCACCGCCATCTT---GTCGATGCCGA 1333683

Qy 250 gAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIl 270
```


Db 1333682 GCGTGTGGCGATGGTCCG-----CAGAT 1333659
Qy 270 eLeuGlyAlaSerArgHisAspGly-AlaTyrSerAspGlyHisPheLeuPheLysProG 290
Db 1333658 CGTCGGC---GAACGCCAAGATGGTTTCGTACTGATCGGCGCTCGAGCGCGCGATGCCGG 1333602
Qy 290 lyGlyThrSerProLeu----- 295
Db 1333601 GCCGATCGACGCCCAATCAGACGAATATGTGGTCTCGGCGTAGACCGGGCTTCGGTCG 1333542
Qy 296 --PheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProp 315
Db 1333541 GGATCTGCCCGCGGGCCCTGGGTGCCATGCAGCCAGAACCGCA----- 1333495
Qy 315 roArgProLeuProArgSerThrPheAlaArgPro 326
Db 1333494 --CGACCTCGGGGCGCGCAACTCGGCGAAGCCG 1333462
RESULT 11
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 5.49e+03 Length: 4411529
Score: 130.50 Matches: 80
Percent Similarity: 36.72% Conservative: 43
Best Local Similarity: 23.88% Mismatches: 120
Query Match: 6.09% Indels: 92
DB: 3 Gaps: 16
US-10-029-020-14_COPY_1_400 (1-400) x US-09-103-840A-1 (1-4411529)
Qy 32 GlyLysAlaProGlnLysSerTyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp 51
Db 1334832 GGTCGGGTCCCAAATCTGCATCAGGGTGCTCAAATCTCTTAC----- 1334788
Qy 52 AlaArgLeuAlaTyrGlySerArgValLysAspIleValProGlnGluAlaGluPhe 71
Db 1334787 -----CACGGCCAAAGCCCGAGAT-----GACTTT 1334761
Qy 72 CysArgThrGlyAlaAsnPheThrLeuArgGluLeuGlyLeuGluValThrProPro 91
Db 1334760 CGCCGGGCAAGTGCAACAACCTTCGGTCGGGAAGCCGGATACGACGTGTT-----CCC 1334707
Qy 92 HisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAla 111
Db 1334706 CGTGGG-----AGAACGGGA-----CGATGTGGTCGTGGTCTCCGTGCC 1334668
Qy 112 GlySer-AspAlaAspMetGluAlaAspThrValLeuSerProGluHisProVal-ArgL 131
Db 1334667 ACCAGCGGACGGGACTTTTAC-----CTCGTCCAGCCGGAATCCCCAGTCGCGGG 1334617
Qy 131 euTlpGlyArgSerThrArgSerGlyArgSerSerCysLeuSer-----SerArgA 148

Db 1334616 CAAAGGCGATGACATCGGCGAACGGCGCAGCGAGCTGCTTCCGACTACCGTTGAGCAGAT 1334557
Qy 148 laAsnSerAsnLeuThrLeu-ThrAspThrGluHisGluAsnThr----- 162
Db 1334556 CGTCGAGGAACATCGCCTTGAACTCGGGCGGAGCCAGCAAAATGCCGTCGCGCCCGCGGTG 1334497
Qy 163 -----GluThrAspHisProGlyGlyLeu 170
Db 1334496 AGAGCAGGCCATACAGGTCGAGGGCAGGGACGCGACGGCGCGGCCCGCGATCAGCA 1334437
Qy 171 GlnAsnHisAlaArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsn 190
Db 1334436 AGCTCGACCCAGCCGACGCGGGGTGCGCCACCTGCAGCAGCGGCGCCACCGCGAAC 1334377
Qy 191 Gln-HisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAs 210
Db 1334376 CAAGGCGCATCAAAACCGCGCTAATCGGTCGCGGCCGCGCT-----CGGCGCGACGC 1334323
Qy 210 nProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaG1 230
Db 1334322 CGCCGAGGACACCGGGCGGACCAACCCGGTC---GGCGAGCCC-----GGCGGCGCAGC 1334272
Qy 230 nGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrAr 250
Db 1334271 CCAGGGTGTATGGGCCCCCGCCCGACAGGCCACCACGGCCATCTT---GTCGATGCCGA 1334215
Qy 250 gAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspI1 270
Db 1334214 GCGTGTGCGCGATGGTCCG-----CAGAT 1334191
Qy 270 eLeuGlyAlaSerArgHisAspGly-AlaTyrSerAspGlyHisPheLeuPheLysProG 290
Db 1334190 CGTCGGC--GAACGCCAAGATGGTTTCGTACTGATCGGCGTCGAGGCGCCGATGCCGG 1334134
Qy 290 lyGlyThrSerProLeu----- 295
Db 1334133 GCCGATCGACGCCCAATCAGACGCAATATTGGTGTCTCGGCGTAGACCCGGCTTCGGTCG 1334074
Qy 296 --PheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProp 315
Db 1334073 GGATCTGCCCGCGGGCCCTGGGGTGCCATGCAGCCAGAACCGCA----- 1334027
Qy 315 roArgProLeuProArgSerThrPheAlaArgPro 326
Db 1334026 --CGACCTCGGGGCGCGCGCAACTCGGCGAAGCCG 1333994
RESULT 12
US-09-252-991A-11748
; Sequence 11748, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11748
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11748
Alignment Scores:
Pred. No.: 0.0321 Length: 1998
Score: 130.00 Matches: 76
Percent Similarity: 33.63% Conservative: 39

Best Local Similarity: 22.22%		Mismatches: 103
Query Match: 6.07%		Indels: 124
DB:		Gaps: 17
US-10-029-020-14_COPY_1_400 (1-400) x US-09-252-991A-11748 (1-1998)		
QY	5 GluArgLysProTyrArgSerLeuThrArgArgArgAlaGluArgArgTyrThrSer	24
Db	913 GAGCGGCATCGCCACCGCGT---GCTCGCCGAAGCGCGCAGACCGTCCGACAGTCCGCG	969
QY	25 SerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSerSerSerGluThr	44
Db	970 TCCGCTGCTGAGCATCGTTCCCGCCGACACCCC-----GTTGACAGGCCGAAC	1017
QY	45 LeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAspIleVal	64
Db	1018 CTACGCGCCGAGCAAGTCCATCGGTTTCATCCGGCCGGGCGACGCGGTGCTGATCCGCTA	1077
QY	65 ProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeuArgGluLeuGly	84
Db	1078 CCA-----GGCCTATCC	1089
QY	85 LeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGln	104
Db	1090 GTACCAGAAAGTTCGGCCA-----GTACCACGG	1116
QY	105 CysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeu---	123
Db	1117 CAAGGT-----GCAGTCGATCTCCCGCGCCAGCGTCTCCTATGC	1155
QY	124 ---SerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArgSerSer	142
Db	1156 CGAGCTTTCAGCATGGTCGGCGGCTACCGG-----GCTCGCCAGGATGGCGAGCA	1209
QY	143 CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThr	162
Db	1210 GCTGTACCGGCTCGGGTAACCTCGA-----	1236
QY	163 GluThrAspHisProGlyGly-----LeuGlnAsnHis	173
Db	1237 -----CGACCAGGCGGTGACCGCTACGGCCAGCGCGTCCGCTGCAGAGCGGCAT	1287
QY	174 AlaArgLeuArgThrProProPro-----ProLeu-----SerHis	185
Db	1288 GCTGCTGACGCGCCGACATCCTCCAGGACACCCGGCGCCTCTACGAATGGGTGCTGGAACC	1347
QY	186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe	205
Db	1348 GCTCTACAGCCTGACCGGCAAACTCTAGGAACGACCCATGGCCTTCTCTCGA-----	1398
QY	206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro	225
Db	1399 -----CGTCTCGCCCTGCGCCTGGGCGG-----CCGCCT	1428
QY	226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle	245
Db	1429 GCGCTGGTGTGTCAGACCGAAGCCACCGA-----ATCGCGCCTGGCCTG	1473
QY	246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu	265
Db	1474 CCTGGCGATGATCGCGGCTACACGGCCACCA-----	1506
QY	266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe	285
Db	1506 -----	1506
QY	286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu	305
Db	1507 ----TACCGGCCTGATGAAGTTCGCGCGGCTTCTCCGTATCGCTCAAGGCGCATCTCCCT	1563
QY	306 ThrSerSerThrValTyrSerProProProArgPro---LeuProArgSerThrPheAla	324
Db	1564 CAAGCAACTGATCCAGACCGCCACCGCCTCGGCTGGGTACCCCGCGCGGTGAA---GCT	1620

QY 325 ArgPro 326
Db 1621 CGACCT 1626

RESULT 13
US-09-252-991A-11618/c
; Sequence 11618, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11618
; LENGTH: 3180
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11618

Alignment Scores:
Pred. No.: 0.0668 Length: 3180
Score: 130.00 Matches: 76
Percent Similarity: 33.63% Conservative: 39
Best Local Similarity: 22.22% Mismatches: 103
Query Match: 6.07% Indels: 124
DB: 4 Gaps: 17

US-10-029-020-14_COPY_1_400 (1-400) x US-09-252-991A-11618 (1-3180)

QY	5 GluArgLysProTyrArgSerLeuThrArgArgArgAlaGluArgArgTyrThrSer	24
Db	2314 GAGCGGCATCGCCACCGCGT---GCTCGCCGAAGCGCGCAGACCGTCCGACAGCTCGCG	2258
QY	25 SerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSerSerSerGluThr	44
Db	2257 TCCGCTGCTGAGCATCGTTCCCGCGCAGACCCC-----GTTGACAGGCCGAAC	2210
QY	45 LeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAspIleVal	64
Db	2209 CTACGCGCGGAGCAAGTCCATCGGTTTCATCCGGCCGGGCGACGCGGTGCTGATCCGCTA	2150
QY	65 ProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeuArgGluLeuGly	84
Db	2149 CCA-----GGCCTATCC	2138
QY	85 LeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGln	104
Db	2137 GTACCAGAAAGTTCGGCCA-----GTACCACGG	2111
QY	105 CysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeu---	123
Db	2110 CAAGGT-----GCAGTCGATCTCCCGCGCAGCGTCTCCTATGC	2072
QY	124 ---SerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArgSerSer	142
Db	2071 CGAGCTTTCAGCATGGTCGGCGGCTACCGGG-----GCTCGCCAGGATGGCGAGCA	2018
QY	143 CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThr	162
Db	2017 GCTGTACCGGCTCGGGTAACCTCGA-----	1991
QY	163 GluThrAspHisProGlyGly-----LeuGlnAsnHis	173
Db	1990 -----CGACCAGGCGGTGACCGCCTACGGCCAGCCGCGTCCGCTGCAGAGCGGCAT	1940


```
Db 2019 CATCCAGCACAGCACCACTTCTCTTCTGGGATCCCCACAGGTTTCCAGCCCCAAGGAT 2078
QY 252 euGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuG 272
Db 2079 TGGGCCCCAGCCCC-----AGCCCCATCCTCAGCCCCATCCTTCAACAGCGT---TTGG 2129
QY 272 lyAlaSerArgHisAspGlyAlaTyr---SerAspGlyHisPhe-LeuPheLysProGly 290
Db 2130 GCCTCAGCCCCCACAGCAGCCCCCTTCCACTCCAGCATCCACATCTCTTCCACCCAGGC 2189
QY 291 GlyThrSerProLeu-----PheCys----- 297
Db 2190 CCCAGGACTCCTACCCCCACAATCCCCCTACCCCTATGCCCTCAGCCCTGGGGTCCTGGG 2249
QY 298 -----ThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProPro 315
Db 2250 GCAGCCGCCACCCCCCTACACACCCAGCTCTACCCAGGTCCCGCTCAAGACCCCTCTGCC 2309
QY 316 ArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLys 335
Db 2310 AGCCCACTCAGGGGCTCTGCCTTTCCCCAGCCCTGGGGCCCCCTCAGCCCTCCCATCCCC 2369
QY 336 TyrCysAsnTyrLysCys 341
Db 2370 ACTGGCATATGGTCCTGC 2387
```

Search completed: August 14, 2004, 19:39:29
Job time : 5373.34 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 7074.72 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400
Perfect score: 2143
Sequence: 1 MDVKKPKYRSLTRRDAER.....EITEDTASSWPVTDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spol/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN_1_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29:	em_vi:*
30:	em_htg_hum:*
31:	em_htg_inv:*
32:	em_htg_other:*
33:	em_htg_mus:*
34:	em_htg_pln:*
35:	em_htg_rod:*
36:	em_htg_mam:*
37:	em_htg_vrt:*
38:	em_sy:*
39:	em_htgo_hum:*
40:	em_htgo_mus:*
41:	em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2143	100.0	8354	6	AX556500	Sequence
2	2138	99.8	1680	6	AR270935	Sequence
3	2138	99.8	3111	6	BD270887	Method of
4	2138	99.8	3111	6	AR270934	Sequence
5	2138	99.8	3111	6	BD140153	Utilizati
6	2138	99.8	3111	9	AF009227	Homo sapi
7	2098	97.9	8438	6	AX675551	Sequence
8	2082	97.2	8645	6	AX600210	Sequence
9	2050	95.7	8585	10	AB025413	Mus muscu
10	1971	92.0	9722	10	AF059485	Mus muscu
11	1495.5	69.8	9264	5	AB026980	Danio rer
12	1285	60.0	2387	6	AR270936	Sequence
13	994	46.4	8964	10	AB025412	Mus muscu
14	987	46.1	1476	9	AF195420	Homo sapi
15	987	46.1	8645	6	AX662355	Sequence
16	987	46.1	8675	6	AX662353	Sequence
17	920	42.9	8689	6	AX250067	Sequence
18	920	42.9	8689	10	AF086607	Rattus no
19	917	42.8	2496	5	GGA245711	Gallus ga
20	917	42.8	2496	6	AX250064	Sequence
21	916	42.7	8797	6	AX250063	Sequence
22	916	42.7	8797	6	AX250066	Sequence
23	916	42.7	8797	10	AB025411	Mus muscu
24	897	41.9	8409	5	GGA279031	Gallus ga
25	897	41.9	8409	6	AX250068	Sequence
26	871	40.6	8575	6	AX921803	Sequence
27	839	39.2	9729	6	AX250013	Sequence
28	839	39.2	9826	6	AX250008	Sequence
29	779	36.4	8118	5	GGA238613	Gallus ga
30	736	34.3	8373	10	AB025410	Mus muscu
31	734.5	34.3	8297	9	AF100772	Homo sapi
32	653.5	30.5	8473	6	AX662357	Sequence
33	653.5	30.5	8487	6	AX662359	Sequence
34	534	24.9	1755	6	AX050264	Sequence
35	508	23.7	1727	6	AX714026	Sequence
36	508	23.7	1727	9	AK056053	Homo sapi
37	506.5	23.6	7713	9	HSM808325	Homo sapi
38	487.5	22.7	157677	2	AP002470	Homo sapi
39	487.5	22.7	166063	9	AP002515	Homo sapi
40	487.5	22.7	168308	2	AC084775	Homo sapi
41	476	22.2	87834	9	AP002957	Homo sapi
42	476	22.2	159946	2	AC079194	Homo sapi
43	458.5	21.4	93474	2	AC136801	Rattus no
44	458.5	21.4	173756	2	AC134948	Rattus no
45	458.5	21.4	245722	2	AC120288	Rattus no

ALIGNMENTS

AX556500 LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M., Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E., Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L., Smithson,G., Li,L. and Ji,W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source 1..8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.69e-101 Length: 8354
Score: 2143.00 Matches: 400
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x AX556500 (1-8354)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
DB 35 ATGGACGTGAAGGAGAGGAAGCCCTACCGCTCGCTGACCCCGCGCGCGACGCCGAGCGC 94
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 95 CGCTACACAGACTCGTCCGGGACAGCGAGGAGGGCAAAGCCCCCGAGAAATCGTACAGC 154
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 155 TCCAGCGAGACCCCTGAAGGCCTACGACGAGACGCCCGCTAGCCTATGGCAGCCCGCTC 214
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
DB 215 AAGGACATTGTGCGCGAGGAGCGGCGGAGGAATCTGCCGCACAGGTGCCAACTTCACCCCTG 274
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
DB 275 CGGAGCTGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCTGTACCGGACAGACATT 334
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
DB 335 GGCCTGCCCAATGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 394
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
DB 395 ACGGTGCTGTCCCTGAGCACCCCGTGTGTGGGGCGGAGCACACGGTCAGGGCGC 454
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB 455 AGCTCCTGCTGTCCAGCCGGGCCAATTCGAATCTCACACTCACCGACACCGAGCATGAA 514
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
DB 515 AACACTGAGACTGATCATCCGGCGGCTCGAGAACCACCGCGGCTCCGACGCCGCCG 574
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200

Db 575 CCGCGGCTCTCGACGCCCAACACCCCAACACGACACCGCGGCTCCATTAACTCCCTG 634
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 635 AACCGGGCAACTTTCACGCCGAGGAGCAACCCCAAGCCCGGCCCAACGACCTCGCTC 694
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 695 TCCGGAGAGCCCCCTGCCGGCGGCCAGGAGCCTGCCCAAGCCAGGAGAACTGGCTG 754
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 755 CTCACAGCAACATCCCCCTGGAGACCAAGAACCTTAGGCAAGAGCCATTCTCTAGGACA 814
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 815 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGCTTAC 874
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 875 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCATCA 934
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 935 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 994
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 995 AGCACCTTCGCCCGCGCCCTTAACTCAAGAACCCCTCAAGTACTGTAACTGGAAG 1054
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1055 TGCGCAGCCCTGAGCGCCATCGTCACTCTCAGCCACTCTGGTCACTCTGCTGCATAC 1114
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGlyGlnMetTyr 380
Db 1115 GTGGCCATGCACCTGTTTGGCCCTAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1174
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1175 GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGTCCAAACCGACGTCTCCCTATACCC 1234
RESULT 2
AR270935 LOCUS AR270935 1680 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 3 from patent US 6500941.
ACCESSION AR270935
VERSION AR270935.1 GI:29702177
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1680)
AUTHORS Schaefer,G.M. and Sliwkowski,M.
TITLE Gamma-herregulin
JOURNAL Patent: US 6500941-A 3 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..1680
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 5.11e-102 Length: 1680
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x AR270935 (1-1680)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20

Db 1 ATGGACGTGAAGGAGAGGAGACCTTACCGCTCGCTGACCCGCGCGCCGCGACGCGAGCGC 60

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40

Db 61 CGCTACACCACTCGTCCGCGGACAGCGAGGGCAAGCCCCCGCAGAAATCGTACAGC 120

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60

Db 121 TCCAGCGAGACCCCTGAGGCTAGCACCAGGACGCGCGCTAGCCTATGGCAGCGCGCTC 180

QY 61 LysAspIleValProGlnGluAlaGluGlyPheCysArgThrGlyAlaAsnPheThrLeu 80

Db 181 AAGGACATTGTCCGCGCAGGAGCGCGAGGAATTCTGCGGCACAGGTGCCAACITTCACCCCTG 240

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100

Db 241 CGGAGCTGGGGCTGGAGAGTAACGCCCCCTCAGGGACCCTGTACCGGACAGACATT 300

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120

Db 301 GGCCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTGTATGCCGACATGGAGGCTGAC 360

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140

Db 361 ACGGTGCTGTCCCTGAGCACCCCGTGTGTGGGGCCGGAGCACACGGTCTCAGGGCGC 420

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160

Db 421 AGCTCTCGCTGTCCAGCGGGCCAAATTCCAATCTCACACTCACCGACACCGAGCATGAA 480

QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180

Db 481 AACACTGAGACTGATCATCCGGCGGCTGAGAACCCAGCGCGGCTCCGACGCGCGCG 540

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200

Db 541 CCGCGCTCTCGACGCCCCACACCCCCAACACGACACACCGCGGCTCCATTAACCTCCCTG 600

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220

Db 601 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCGCGCGCCCCACGACCACTCGCTC 660

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240

Db 661 TCCGAGAGCCCCCTGCGCGGCGCCAGGAGCCTGCCCCACGCGCGAGGAACTGGCTG 720

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260

Db 721 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCATTCCTAGGGACA 780

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280

Db 781 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGCGCCTCCCGCATGATGGGGCTTAC 840

QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300

Db 841 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCACATCA 900

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320

Db 901 CCAGGTACCCCACTGACGTCCAGCACAGTGACTCTCTCCGCGCGGACCCCTGCCCGCG 960

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340

Db 961 AGCACCTTCGCGCGCGGCGCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACCTGGAG 1020

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360

Db 1021 TCGCAGCGCCTGAGCGCCATCGTCTCTCAGCCACTCTGGTCTATCTGTGGCATACTTT 1080

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1081 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1140

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1141 GAGATCAGGAGGACACAGCCAGAGTTGGCCTGTGCCAACCGACGTCCTCCCTATACCCC 1200

RESULT 3

BD270887

LOCUS

DEFINITION

3111 bp DNA linear PAT 17-JUL-2003

Method of promoting the proliferation of inner ear hairy cells

using ligand of HER2 receptor and/or HER3 receptor.

BD270887

ACCESSION

BD270887

VERSION

BD270887.1 GI:33080655

KEYWORDS

JP 2002529425-A/6.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3111)

Gao,W.Q.

Method of promoting the proliferation of inner ear hairy cells

using ligand of HER2 receptor and/or HER3 receptor

Patent: JP 2002529425-A 6 10-SEP-2002;

GENENTECH INC

JOURNAL

OS Homo sapiens (human)

PN JP 2002529425-A/6

PD 10-SEP-2002

PF 28-OCT-1999 JP 2000580655

PR 07-NOV-1998 US 60/107522

PI WEI QIANG GAO

PC A61K45/00,A61K38/00,A61K39/395,A61P27/16,C12N5/06//(C12N5/06,

PC C12R1:91),

PC A61K37/02,C12N5/00,(C12N5/00,C12R1:91)

CC Method of promoting the proliferation of

inner ear hairy cells

using

CC ligand of HER2 receptor and/or HER3 receptor

FH Key Location/Qualifiers

FT source 1..3111

FT /organism='Homo sapiens (human)'

FEATURES

source

1..3111

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.02e-101 Length: 3111

Score: 2138.00 Matches: 399

Percent Similarity: 99.75% Conservative: 0

Best Local Similarity: 99.75% Mismatches: 1

Query Match: 99.77% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x BD270887 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20

Db 334 ATGGACGTGAAGGAGAGGAGGAGCCTTACCGCTCGCTGACCCGCGCGCGCGCGCGC 393

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40

Db 394 CGCTACACCACTCGTCCGCGGACAGCGAGGAGGGCAAGCCCCGAGAAATCGTACAGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60

Db 454 TCCAGCGAGACCCCTGAAGGCTACGACCAAGGACCGCGCGCTAGCCTATGGCAGCGCGC 513

QY 61 LysAspIleValProGlnGluAlaGluGlyPheCysArgThrGlyAlaAsnPheThrLeu 80

Db 514 AAGGACATTGTCCGCGCAGGAGCGCGAGGAATTCTGCCGCGACAGGTGCCAACTTCACCCCTG 573

QY	81	ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	574	CGGAGCTGGGCTGGAAGAAAGTAACGCCCTCACGGACCTGTACCGACAGACATT	633
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	634	GGCTCCCCCACTGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGTGAC	693
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	694	ACGGTGCTGTCCCTGAGCACCCCGTGGTCTGTGGGGCCGGAGCACACGGTCAGGGCGC	753
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	754	AGCTCCTGCTGTCCAGCCGGGCCAATTCGAATCTCACACTACCGACACCGAGCATGAA	813
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	814	AACACTGAGACTGATCATCCGGCGGCTGCAGAACCCACGCGGCTCCGGACGCGCGC	873
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu	200
Db	874	CCGCCGCTTCGCACGCCCACACCCCAACAGCACCCAGCCCTCCATTAACCTCCCTG	933
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProAlaProThrAspHisSerLeu	220
Db	934	AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGCCCGACCACTCGCTC	993
QY	221	SerGlyGluProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	994	TCCGGAGAGCCCTGCGGCGGCGCCAGGAGCTGCCACGCCAGGAGAACTGGCTG	1053
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr	260
Db	1054	CTCAACAGCAACATCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCCTAGGGACA	1113
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	1114	TTGCAGGACAACCTCATTGAGATGGACATTCTCGGCGCTCCCGCCATGATGGGCTTAC	1173
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	1174	AGTGACGGGCATTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCAATCA	1233
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg	320
Db	1234	CCAGGGTACCCACTGACGTCCAGCACAGTGTACTTCTCCGCCCGACCCCTGCCCGC	1293
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys	340
Db	1294	AGCACCTTCGCCCGCGCGCTTAACTCAAGAAAGCCCTCCAAGTACTGTAACTGGAAG	1353
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1354	TGCGCAGCCCTGAGCGCATCGTCACTCAGCCACTCTGGTCACTCCTGCTGGCATACTT	1413
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1414	GTGGCCATGCACCTGTTGGCTTAAACTGACCTGACCTGAGCCGATGGAGGGGAGATGAT	1473
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	1474	GAGATCAGGAGGAGACACAGCAGCAGTGGCTGTGGCTGTGCCAACCGACGTCTCCCTATACCCC	1533

RESULT 4	AR270934	3111 bp	DNA	linear	PAT 10-APR-2003
LOCUS	Sequence 1	from patent US 6500941.			
DEFINITION	AR270934				
ACCESSION	AR270934				
VERSION	AR270934.1	GI:29702176			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				

REFERENCE 1 (bases 1 to 3111)
AUTHORS Schaefer,G.M. and Sliwowski,M.
TITLE Gamma-heregulin
JOURNAL Patent: US 6500941-A 1 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..3111
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores: 1.02e-101 Length: 3111
Pred. No.: 2138.00 Matches: 399
Score: 2138.00
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AR270934 (1-3111)

QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	334	ATGGACGTGAAGAGAGAGAGCCTTACCGCTCGCTGACCCGGCGCCGACGCCGAGCGC	393
QY	21	ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	394	CGCTACACCACTCGTCCCGGACAGCGAGAGGGCAAGCCCCGCGAGAAATCGTACAGC	453
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	454	TCCAGCGAGACCTGAAGCCCTACGACCGAGACGCCCTAGCCTATGGCAGCCGCGTC	513
QY	61	LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu	80
Db	514	AAGGACATTGTGCCGAGGAGGCGCGAGAAATCTGCCGCACAGGTGCCAACTTCACCTG	573
QY	81	ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	574	CGGAGCTGGGCTGGAAGAAGTAACGCCCTCACGGACCCCTGTACCGACAGACATT	633
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	634	GGCTCCCCCACTCGGGCTACTCCATGGGGGTGGCTCTGTATGCCGACATGGAGGTCAC	693
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	694	ACGGTGCTGTCCCTGAGCACCCCGTGGCTGTGGGGCCGGAGCACACCGTCAGGGCGC	753
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	754	AGCTCCTGCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA	813
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	814	AACACTGAGACTGATCATCCGGCGGCTGCAGAACCCACCGCGGCTCCGGACGCCCGCG	873
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu	200
Db	874	CCGCCGCTCTCGCACGCCCAACCCCCAACAGCACCGCGGCTCCATTAACTCCCTG	933
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	934	AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCGGCCCCCACCGACCACTCGCTC	993
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	994	TCCGGAGAGCCCTGCGGCGGCGCCAGGAGCTGCCACGCCAGGAGAACTGGCTG	1053
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr	260
Db	1054	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCCTAGGGACA	1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
| | | | |
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGGGCTCCCGCCATGATGGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
| | | | |
Db 1174 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
| | | | |
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
| | | | |
Db 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
| | | | |
Db 1354 TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTT 1413

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
| | | | |
Db 1414 GTGGCCATGCACCTGTTGGCCTAAACTGGACCTGCAGCCGATGGAGGGGCAGATGTAT 1473

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
| | | | |
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCC 1533

RESULT 5
BD140153
LOCUS BD140153 3111 bp DNA linear PAT 18-SEP-2002
DEFINITION Utilization of heregulin as epithelial cell growth factor.
ACCESSION BD140153

VERSION BD140153.1 GI:23235098
KEYWORDS JP 2002509076-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 3111)
TITLE Sliwkowski, M. and Kern, J.A.
JOURNAL Utilization of heregulin as epithelial cell growth factor
Patent: JP 2002509076-A 6 26-MAR-2002;
GENENTECH INC, UNIVERSITY OF IOWA RESEARCH FOUNDATION

COMMENT OS Homo sapiens (human)
PN JP 2002509076-A/6
PD 26-MAR-2002
PF 03-FEB-1999 JP 2000530226
PR 04-FEB-1998 US 09/020598
PI MARK SLIWKOWSKI, JEFFREY A KERN
PC A61K38/00, A61K35/12, A61K39/395, A61K39/395, A61P11/00, A61P43/00,
PC C07K14/485,
PC C07K16/22, C12N15/09, C12P21/08, A61K37/02, C12N15/00 CC
Utilization of heregulin as epithelial cell growth factor FH Key
Location/Qualifiers
FT source 1..3111
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1..3111
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.02e-101 Length: 3111
Score: 2138.00 Matches: 399
Percent Similarity: 99.75% Conservative: 0
Best Local Similarity: 99.75% Mismatches: 1
Query Match: 99.77% Indels: 0
DB: 6 Gaps: 0

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
| | | | |
Db 334 ATGGACGTGAAGAGAGAGAGCTTACCGCTCGCTGACCCGGCGCGACGCCGAGCGC 393

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
| | | | |
Db 394 CGCTACACCACTCGTCCGGGACAGCGAGAGGGGAAAGCCCCCGCAGAAATCGTACAGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
| | | | |
Db 454 TCCAGCGAGACCTGAAGGCCTACGACCGACGCCCGCTAGCCTATGGCAGCGCGCTC 513

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
| | | | |
Db 514 AAGGACATTTGCGCGCAGGAGGCCGAGGAATTTGCCGCACAGGTGCCAACTTACCCCTG 573

QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
| | | | |
Db 574 CGGAGCTGGGCTGGAAGAAGTAACGCCCTTACGGGACCTGTACCGACAGACATT 633

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
| | | | |
Db 634 GGCCTCCCCCACTGGGCTACTCCATGGGGGCTGGCTCTGTATGCCGACATGGAGGCTGAC 693

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
| | | | |
Db 694 ACGTGTCTTCCCTGAGCACCCCGTGTCTGTGGGCGGAGCACACGGTCAGGGCGC 753

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
| | | | |
Db 754 AGTCTCTGCTGTCCAGCGGGCCAAATTCGAATCTCACACTCACCGACACCGAGCATGAA 813

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
| | | | |
Db 814 AACACTGAGACTGATCATCCGGCGGCTGTCAGAACCCACCGCGGCTCCGACGCCCGCG 873

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
| | | | |
Db 874 CCGCGGCTCTCGCAGCCCAACACCCCAACAGCACCCACCGCGGCTCCATTAACTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
| | | | |
Db 934 AACCGGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCGCCAGCAACCACTCGCTC 993

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
| | | | |
Db 994 TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACCGCCAGGAGAACTGGCTG 1053

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
| | | | |
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCCAGAAACCTTAGGCAAGCAGCATTCCTAGGACA 1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
| | | | |
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGCGCCTCCCGCATGATGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
| | | | |
Db 1174 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGCACCTTCCCGCTCTTCTGCACCATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
| | | | |
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
| | | | |
Db 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
| | | | |
Db 1354 TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTT 1413

QY 381 GluileThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGAGTTGGCCTGTGCCAACCGACGTCTCCCTATACCCC 1533

RESULT 7
LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO02055704.
ACCESSION AX675551
VERSION AX675551.1 GI:29333352
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
Stone,D.J.

TITLE Proteins, polynucleotides encoding them and methods of using the
same

JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)

FEATURES
source Location/Qualifiers
1..8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 3.61e-99 Length: 8438
Score: 2098.00 Matches: 398
Percent Similarity: 98.76% Conservative: 0
Best Local Similarity: 98.76% Mismatches: 2
Query Match: 97.90% Indels: 3
DB: 6 Gaps: 1

US-10-029-020-14_COPY_1_400 (1-400) x AX675551 (1-8438)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
Db 4 ARGACGTGAAGGAGAGGAGCGCTTACCGCTCGCTGACCCGGCGCGGACGCGGAGCGC 63

QY 21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 64 CGCTACACCACTCGTCCGGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC 123

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 124 TCCAGCGAGACCCCTGAAGGCCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCCGCTC 183

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 184 AAGGACATTGTGCCGCGAGGAGCCGAGGAATTCTGCCGACAGGTGCCAACTTCACCCCTG 243

QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 244 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCTCACGGACCCCTGTACCGGACAGACATT 303

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 304 GGCCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGSGTGAC 363

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 364 ACGGTGCTGTCCCTGAGCACCCCGTGCCTCTGTGGGGCGGAGCACACGGTCAGGGCGC 423

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnSerLeuThrLeuThrAspThrGluHisGlu 160
Db 424 AGCTCTCTGCTGTCCAGCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 483

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 484 AACACTGAGACTGATCATCCGGGGCGCTGCAGAACACACGCGGCTCCGGACGCGCGG 543

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 544 CGCCCGCTCTCGCAGCCCAACACCCCAACACGACACACGCGGCTCCATTAACTCCCTG 603

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 604 AACCGGGGCACTTCAGCCGAGGAGCAACCCAGCCCGCCCCACGGACCACTCGCTC 663

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 664 TCCGGAGAGCCCTGCCGGCGGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG 723

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 724 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGACA 783

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 784 TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGGCGCTCCCGCCATGATGGGCTTAC 843

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 844 AGTGACGGGCACCTCTCTTCAAGCCTGGAGGCACCTCCCCCGCTTCTTGACCACTCA 903

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 904 CCAGGGTACCCACTGACGTCACGACAGTGTACTCTCTCCCGCCGACCCCTGCCCGC 963

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 964 AGCACTTCGCCTGGCGGCCCTTAACCTCAAGAAGCCCTCCAAGTACTGTAACCTGAAG 1023

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1024 TGCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTTCATCTCTGGCATACTTT 1083

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
Db 1084 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGACG 1143

QY 380 ---TyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
Db 1144 GATTTATGAGATCACGGAGGACACAGCCAGCAGTTGGCTGTGCCAACCGACGTCTCCT 1203

QY 398 uTyrPro 400
Db 1204 ATACCCC 1210

RESULT 8
AX600210
LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Walia,N.K.

TITLE Proteins associated with cell growth, differentiation, and death

DGAYS DGHFLFKPGGTSPLFCTTSPGYPLTSSTVYSPPPRPLPRSTFSRPAFLKKPS
KYCNMKCAALSAILISATLVILLAYFAMVHLFGLNWHLQPMEGQMOMYEITEDTASSW
PVPTDSLYPSGGTGLETPDRKGGAEGKSPSSLFPEDSFDSIEDVGGRRASQKIPP
GTFWRSQVFIIDHPVHLKFNVSLGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQE
ARSLLEGQRPQSGPVPSPSSHETGFIQYLDSGIWHLAFYNDGKESEVVFLTTAIESVD
NCPNSCNYGNDICISGTCFLGFLGPDGCRASCPVLCSNGQYMKRCLCHSGWKGA
CDVPTNOCIDVACSHGTCIMGTICINPGYKGESCEEVDMDPTCSSRGVCVRGEC
SVGWGTTNCETPRATCLDQCSEHGTFPLDPTGLCNCDPSTWTHDCSIEICAADCGGHV
CVGTCRCEDGWMGAACDQACHPRCAEHGTCRDCKCECSPGWNGEHCCTIAHYLDVRV
KEGCPGLCNGNRCTLDLNGWHVCQLGWRGTGCDTSMETGCGDKDNDGDGLVDCMD
PDCCQLPCHVNPLCLGSPDPLDIQETQAPVSQQNLNPFYDRIKFLVGRDSTHSIPG
ENPFGGHACVIRGQVMTSDGTPLVGNISFINNPLFGYTISRQDGSFDLVTNGGISI
ILRFERAPFITQEHTLWLPWDRFFVNETIVMRHENEIPSCLDSNFARPNPVSPSPL
TSFASCAEKGPVIVPEIQALQEEIVTAGCKMRLSYLSRSTPGYKSVLRISLTHPTIPF
NLMKVHLMVAVEGRLFRKWFAAAPDLISYFIWDKTDVYNQVFGFSEAFVSUVGYEYES
CPDILLWEKRTAVLQGEYIDASKLGWSLDKHALNIQSGILHKGNGENQFVSQQPPV
IGSTMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVDFNYIRRIFFSGNV
TNILEMRNKDFRHSHPAHKYLATDPMGSAVFLSDTNSRRVPFKVSTTVVKDLVKN
EVVAGTGDDQLPFDDTRCGDGGKATEATLTNPRGITVDFGLIYFVDGTMIRRDQNG
LIISTLLSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYLDNNVVLQISE
NHQVRIVAGRPMHCQVPGIDHFLSKVAIHATLESATALAVSHNGVLVYIAETDEKLN
RIRQVTSAGEISLVAGAPSGCDCKNDANDCFSGDDGYAKDAKLTNPSLAVCAGDEL
YVADLGNIRIRFIRKNKPFNLTONMYELSSPIDOELYFDTSGKHLTYQSLPTGDYLY
NFTYTGDDDIITHITDNNGMVNRDSTGMPLWLVPDQVYVWVTMGTSNALSRSVTTQ
GHELAMNTYHGNSGLLATKSNENGWTFYEDYSFGRLTNVTEPTQVVSFRSDTDSSV
HVQVETSKDDVTITNLSAGAPYTLQDQVRNSYYIIGADQSLRLLLLANGMEVALQT
EPHLLAGTVNPTVGRNVTPLIDNGLNLVEWRQKEQARGQVTVFGRRLRVHNRNLLS
LDPDRVTRTEKIYDDHRKFTLRILYDQAGRPSLSWSPSSRLNGVNVITYSPGGHIAQIOR
GIMSERMEYDQAGRITSRIFADGKWSYTYLEKSMVLHLHSQRYIFEFDKNDRLSV
TMENVARQTLLETIRSVGYRNIYQPPEGNASVIOFTEDGHLHTFYLGTGRRVIVKY
GKLSKLAETLYDTTKVSFTYDETAGMLKTVNLONEGFTCTIRYQIGPLIDRIQIFPT
EEGMVNARFDYNDNSFRVTSQAVINETPLPIDLYRVDVSGKTEQFGKFGVIVYDI
NQIITAVMTHTKHFDAYGRNKEVQYEIFRSLMYMTVQYDNMGRVVKELKVGPYAN
TTRYSEYDADGLOQTVSINDKPLWYSYDLNGLNLHLISPGNSARLTPLRYDLDRIT
PLGDVQYKMDDEDFLQRGGDVFEYNSAGLLIKAYNRASGWSVRYRVDGLRRVSSKS
SHSHLQFFYADLTNPTKVTHLYNHSSEITSLYYDLOQHLFAMELSSGDEFYIACDN
IGTPLAVFSGTGLMIKQILYATGEIYMDTNPNFQIIIGYHGGLYDPLTLKLVHMGRRD
YDVLAGRWTSPDHELWKRLSNSIVPFHLYMFKNNPISNSQDIKCFMTDVNSWLTF
ERFDLHNVIIPGYPKPDTDAMEPPSYELVHTQMKTOEWDNSKSLGVQCEVQKQLKAFVTL
ERFDQLYGSTITSCQQAPETKFASGSGIFGKGVKFAKDKGRVTTDIIISVANEDGRRRI
AAILNNAHYLENLHFTIDGVDTHYFVKPGPSEGDLAILGLSGGRTTLENGVNVTVSQI
NTMLSGRTRRYTDIQLQYRALCLNTRYGTTVDEEKRVLRLARQRAVRQAWAREQORL
REGEGLRAWTDGKEKQVLTNTRGVQYDGFVFTSVQYPELSDSANNIHFMRQSEMGR
R"

ORIGIN

Alignment Scores:
Pred. No.: 1.11e-96 Length: 8585
Score: 2050.00 Matches: 382
Percent Similarity: 96.52% Conservative: 6
Best Local Similarity: 95.02% Mismatches: 12
Query Match: 95.66% Indels: 2
DB: 10 Gaps: 1

US-10-029-020-14_COPY_1_400 (1-400) x AB025413 (1-8585)

QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg	20
Db	188	ATGGACGTGAAGGAGAGAGAGCCCTATCGCTCGTGACCGCGCTCGCGATGCTGAGCGC	247
QY	21	ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	248	CGCTACACCACTCATCGGCAGACAGCGAGGAGGCAAGGSCCCACAGAAGTCCCTACAGC	307
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	308	TCCAGTGAGACCCCTGAAGGCATATGACCAAGATGCTCGCCTAGCCTACGGCAGCCGCGTC	367
QY	61	LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu	80
Db	368	AAGGACATGGTACCACAGGAGCCGAGGAGTTCTTGCCGCACAGGCACATAATTTCACTCT	427
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100

Db	428	CGTGAAC TGGACTGGGAGAGATGACGCCCCCTCATGGACTCTCTACAGGACAGACATC	487
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	488	GGCCTCCACACTGTGGCTATTCCATGGGGGCCAGCTCTGATGCAGACTTGGAAAGCAGAC	547
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	548	ACTGTGCTATCCCTGAACACCCAGTGCGCCTGTGGGCGCGGAGCACACGGTCAGGGCGC	607
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	608	AGTTCTCTGCTGTCCAGCCGGGCCAACTCCAACCTCACACTCACGGACACAGCATGAG	667
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	668	AACACAGAAACCGATCATCCAGCAGCCTGCAGAAACCCCTCGGCTCCGGACGCCCTCCT	727
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu	200
Db	728	CCACCACTGCCCCAGCCCATACCCCAACAGCACCAACCGCGCCTCCATCAACTCCTTG	787
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	788	AACAGGGGCAACTTCACCCCAAGGAGCAACCCAGCCAGCCACCCACAGACCACTCACTC	847
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	848	TCTGGGAGCCCCAGCAGGCGCGCCAGGAGCCAAACCCATGCCAGGACAACTGGCTA	907
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
Db	908	CTCAACAGTAACATCCCACTGGAAACACAGAAACCTTAGGCAAGCAGCCATTCTTAGGACA	967
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	968	TTGCAGGACAACTCATTTAGATGAGACATTCTCAGCGCCTCCCGCCATGATGGGCTTAC	1027
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	1028	AGTGACGGGCACCTTCTCTTCAAGCCCGGAGGCACCTCCCCGCTCTTCTGCACCACATCC	1087
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320
Db	1088	CCAGGTTACCCCTTAACGTTCTAGCACCGTGTACTCGCCCCCACCCCGGCCCTGCCCGC	1147
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340
Db	1148	AGCACCTTCTCCCGACCGAGCCTTTAACCTCAAGAAAGCCTTCCAAGTACTGCAACTGGAAG	1207
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1208	TGTGACGGCTGAGCGCCATCCTCATCTCAGCTACGCTCGTCATCTCTGCTGCATACCTT	1267
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGly-----Gln	378
Db	1268	GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGCAG	1327
QY	379	MetTyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeu	398
Db	1328	ATGTATGAGATCACGGAGGACACAGCCAGCAGCTGGCTGTGCCAACGGACGCTCTCCCTG	1387
QY	399	TyrPro 400	
Db	1388	TATCCC 1393	

RESULT 10
AF059485
LOCUS
DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614

KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 9722)
AUTHORS
Wang,X.Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P.,
Zinszner,H. and Ron,D.
TITLE
Identification of novel stress-induced genes downstream of chop
JOURNAL
EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE
98315054
PUBMED
9649432
REFERENCE
2 (bases 1 to 9722)
AUTHORS
Wang,X.-Z. and Ron,D.
TITLE
Direct Submission
JOURNAL
Submitted (14-APR-1998) Skirball Institute, New York University
Medical Center, 550 First Ave., New York, NY 10016, USA
FEATURES
source
Location/Qualifiers
1. .9722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1. .9722
/gene="Doc4"
/note="downstream of CHOP 4"
583. .9060
/gene="Doc4"
/note="similar to Drosophila melanogaster tenm/odz and
human gamma-herregulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translation="MDVKERKPYRSLTRRRDAERRYTSSADSEEGKGPQKSYSSSET
LKAYDQARLAYGSRVKDMVPQEAHEFCRTGTNFTLRELGLCEMTPPHGTLYRTDIGL
PHCGYSGMAGSSADLEADTVLSPHPVRLWGRSTRSGRSSCLSSRANSNLTLDTEHE
NTEGAPLHCSSASSTPIEQSPSPPPANESQRRLLGNGVAQTPDSDSEEFVFN
SFLVKSASGLVAANDHPSSLQNHPRLTTPPPPLPHAHTPNQHAASINSLNRGNFT
PRNSFPAPTDHSLSGEPPAGSAQEPHTAQNWNVLSKIPVETRNLGKQPFGLTWQDN
LIEMDIFASRRDGYSDGFFFKPGTSPLECTSPGYPLTSTVYSPPPRLPRST
FSRPNLKKPSKYNWKAALSAILISATLVILLAYFVAMHLFGLNWHLQPMEGQM
MYEITEDASSWPVPTDVSYPGSGTGLETPDRKGKGAAGKPSLFPEDSFIDSGEI
DVGRRASQKIPPGTFWRQVFIHPVHLKFNVSLSGKAAVLGVIYGRKGLPPSHTQLDFV
ELLDRRLLTQEARSLGEPQQRSGVPVPPSSHETGLFIQYLDGSIWHLAFYNDGKESEV
VSFLTATIESVDNCPSNYNGDCISGTCHEFLGFLGPDCCGRASCPVLCSGNGQYMKG
RCLCHSGWKGAECDVPTNQCIDVACSSHGTCIMGTICINPGYKGESCEEVDCMDPTCS
SRGVCVRGECSCSVGWGNTCEPPTATCLDQCSHGHTFLPDTGLCNDPSPWTHDCSI
EICAACDGGHGVCGGTCRCECDGWMGAACDQACRCAEHTCRDCKCECTPGWNGE
HCTIEGCGLCNGRCRTLDLNGWHCVQQLGWRGTGCDTSMETGCGDGKONDGDGLVD
CMDPDCLQLCHVNPLCLGSPDLIIQETQAPVSQQNLNSFYDIKFLVGRDSTHS
IPGENPFDGGHACVIRGQVMTSDGTPLVGNISFINNPLFGYTI SRQDGSFDLVNTGG
ISIIILFERAPFITQETHLWLPWDRFFVMTIVMRHEENEI PSRDLSNFARPNPVVSP
SPLTSFASCAEKGPVPEIQALQEEIVIAGCKMRLSYLSRSTPGYKSVVRI SLTHTPT
IPFNLKMKVHLMVAVEGRLFERKWFAPADLSYFIWDTDVYNQKVFGLSEAFVSUYGE
YESCPDLILWEKRTAVLQGYEIDASKLGGWSLDKHHALNIQSGILHKGNGENQFVFSQ
PPVIGSIMENSHSPAHKYI LATDPMGSAFVFLSDTNRRRVFKVKSTTVKVDLVKNSEVVA
GTVNTNILEMHSPTAHKYIATDPMGSAFVFLSDTNRRRVFKVKSTTVKVDLVKNSEVVA
GTGDQCLPDDTDCGDDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRVDDQNGIIST
LLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYVLDDNVVLQISENHQV
RIVAGRPMHCQVPQIDQDQLLSKVAIHATLESATLAVSHNGVLYIAETDEKKINRIYD
VITTAGRPMHCQVPQIDQDQLLSKVAIHATLESATLAVSHNGVLYIAETDEKKINRIYD
LTNIRIRIRIRKPKPFLNTQNNYELSSPIDQELYLFDTSKHLTYQSLPTGDLYLNFTY
TGDDGI THITDNGNMVNRDSTGMPLMLVVPDQVYVWVTGNTSNALRSVTTQGHGL
AMTYHGNSGLLTKSNENGWTTTFEYDSFGLTNTVFTPTQVSSFRSDTSSSVTQV
ETSSKDDVTITHLSGSAFYTL LQDQVNSYYIGADGSLRLLALLANGMEVALQTEPHL
LAGTNNVTGKRNVTLPIDNGLNLVWRQKQEQARQGVTVFGRRLRWHNRNLLSLDLD
RVTRTEKIYDHRKFTLRILYDQAGRPSFWSRSLNGVNTYSPGGHIAGIQRIQIMS
ERMEYDQAGRTISRI FADGKMWSYTYLEKSMVLHLHSORQYIFEFKDNRLSSVTMPN
VARQTLTETIRSVGYIRNYQPPEGNASV IQDFTDGHLLTFLYLTGTRVYIKYKGLS
KIAETLYDTTKVSFTYD ETAGMLKTIVNLONEGFTCTIRYROI GPLDIQRIFRFTTEEGM

Db	1243	TTAGTTAAAGTGGCTCCGCCAGCCTGGGGTCCGAGCAAACGATCATCCAGCAGCCTG	1302
QY	171	GlnAsnHisAlaArgLeuArgThrProProProLeuSerHisAlaHisThrProAsn	190
Db	1303	CAGAACCACTCGGCTCCGACGCCTCCTCCACCACTGCCCACTGCCCATACCCCAAC	1362
QY	191	GlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn	210
Db	1363	CAGCATCACGGGCTCCATCAACTCCTTGAACAGGGCAACTTCACCCCAAGGAGCAAC	1422
QY	211	ProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln	230
Db	1423	CCCAGCCCAGCACCCACAGACCACTCACTCTCTGGGAGCCCCAGCAGCGGCCACG	1482
QY	231	GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArg	250
Db	1483	GAGCCAAACCATGCCAGGACAACACTGGGTACTCAACAGTAAATCCCACTGGAACCA	1542
QY	251	AsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIle	270
Db	1543	AACCTAGGAAGCAGCCATTCTTAGGGACATGGCAGGACAACCTCAATTGAGATGGAC	1602
QY	271	LeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysProGly	290
Db	1603	TTCAGCGCTCCCGCGTGATGGGGCTTACAGTGACGGCACCTCTCTTTTCAAGCCCG	1662
QY	291	GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal	310
Db	1663	GGCACCTCCCGCTCTTCTGCACCACATCCCCAGGGTACCCCTTAACGCTTAGCAC	1722
QY	311	TyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu	330
Db	1723	TATTGCCCCACACCCGGCCCCCTGCCCGCAGCACTTTCTCCCGACGACCTTTAAC	1782
QY	331	LysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSer	350
Db	1783	AAGAAGCTTCCAAGTACTGCAACTGGAAGTGTGCAGCGTGACGCGCATCCTCATCT	1842
QY	351	AlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrp	370
Db	1843	GCTACGCTCGTCATCCTGCTCGCATACTTGTGGCCATGCACCTGTTGGCCTAAACT	1902
QY	371	HisLeuGlnProMetGluGly-----GlnMetTyrGluIleThrGluAspThrAlaSer	388
Db	1903	CACCTGCAGCCGATGGAGGGCAGATGCAGATGTATGAGATCAGGAGGACACAGCC	1962
QY	389	SerTrpProValProThrAspValSerLeuTyrPro 400	
Db	1963	AGCTGGCCTGTGCCAACGGACGCTCTCCCTGTATCCC 1998	
RESULT 11			
AB026980			
LOCUS			
DEFINITION			
AB026980			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Danio rerio (zebrafish)			
Danio rerio			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;			
Cypriniformes; Cyprinidae; Danio.			
1 (sites)			
Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.			
Compartmentalized expression of zebrafish ten-m3 and ten-m4,			
homologues of the Drosophila ten(m)/odd Oz gene, in the central			
nervous system			
Mech. Dev. 87 (1-2), 223-227 (1999)			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
2 (bases 1 to 9264)			
Mieda,M.			
Direct Submission			

JOURNAL	Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute, RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hirosawa, Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp, Tel:81-48-467-9713, Fax:81-48-467-9714)
COMMENT	Sequence updated (29-Jun-1999).
FEATURES	Location/Qualifiers
source	1. .9264
	/organism="Danio rerio"
	/mol_type="mRNA"
	/db_xref="taxon:7955"
gene	/dev_stage="embryo"
	1. .9264
	/gene="ten-m4"
CDS	323. .8797
	/gene="ten-m4"
	/note="similar to Drosophila melanogaster tenm/odz and Mus musculus Doc4"
	/codon_start=1
	/product="ten-m4"
	/protein_id="BAA81893.1"
	/db_xref="GI:5307785"
	/translation="MEVKERRPYRSLTSRRDTERRYTSSSADSEDGKINPKSYSSET LKAFQDSRLAYGSRVKDLVHHEADEFSRQGPDFSRLDMAFGDPVPPHMGAYRTEMGL PHRDYSVSVASDADTETDGIMSPHAHVRLWGRSNTKGRSSCLFSRANSNLITDTEH ENTENGPPLHCSSASSPVDSPYPPPSHAANQSRGLGNSGAQAGRDSSEDEFGN SFLVKTGSGNVCTPAATANEGSFQNHSLRTPPLPLFHSHPSQHHTASIGLSRSN YQRSNPSPAPTSSAPNEGPTSAQDSSAQDNWLNNSNPLETNRNIAKQTFLETLQD NFIEMDILATARRDGAYTDGHFLFKPGGTSPLYCTTSPGYPLTSSTVSPPPRLPRN TFSRPAFSLKKPYKHCNWKCAALSAILISVTLVFLLAYFIAMHLFGLNWHLQPVQRQI YQLTEDNTSGLHLPTDLGLPPLGNTGLEFPDRGRDDGKLDGFFPEDSFIDMGEIDVG RKVAQLIPPGIFWRSQVFIDHPMYLKFNVSLSKDALVGIYGRRLPPSHTQFDFELL DRRLLSQGLPGDGPFPQAQRSLVPITSHDTGCIQYMDSGIWHLAVYNDGKETEOV SFLTATIDSIDDPCSNCFNGDCVSGNCHCFPGFRGPDCSRASCPLVCSGNGQYLKGR CMCHSGWKSECDVPTNQCIDITCSGHGTCTVGTCTCNPSYKGCNCEEEVDCDPTCSG RGVCVRGECHCFVCGWGPCCESPRASCMQCSGHGSLADTNTCNDHNWTHDCSTE LCAADCGGHGICVAGSCRCDEGWMGTCEORACHPRCSEHGTCCKGCECSPWNGEH CTIEGCPGLCNGNRCITLGNNGWYCVQALGWRGAGDTSMETACSDGKNDGDLTDC MDPDCCLOASCHTTSLCVGSPDPLDIIQETQISSSLTLQSFYQRIHFLVGRDSTHVI PDVNPFDGIHACVIRGOVVTSDGTPLVGVNISFINPKPAYGTYITRQDGSFDLVNSGV AIGLRFERAPFITQHTLWLPGWFFVMDTIVMRHEVNDIPSCDLSSTFRMPILPA PLTAFAGTCPERGIVVPEIQTLQEEVRIPGTDMRLGLYSRTSGYKSLRLITLTHSTI PFSLMKVHLMVAVEGLFRKWFSAAPNLSYDFVMDKTDVYSQVYGLSEAFVSUGFEY ESCPDLILWEKRTAVLQGYETTASNLGGSVDKHALNIQSGLHKGNGENIFISQOP PVIGSIMGNRRRSISCPSCNGLADGNKLLAPALACGSDGSLYVGDFNYRRIPTTG NVTSVLELSNPAHKYLYATSPVSGWLYLSDTSSRKVFVKVSLYAKDVAKNLELVAG TGDQCLPYDETRCGDGKXAVEATLTNPRGITVDKYGVIFFDVGTMRIRIDQNGHISTL LGFNDLTSARPLSCDSVMDISQVRLEWPTDLAVSPMDNSLYLDNNVVLQISENHQVR IVAGRPMHCQVPLGDHFLVSKIAIHATLESANALAVSHNGLLYIAESDEKKNRVRQV STNGEISLLAGAPSGCDCKNDANDCDYSGDDGAKDAKLNAPSSLAVSPDGELFIADL GNIRIYVRNKAFLNPLNMYEISSPIDDELXLFUDVNASHVFTQSLTGDLYNFTYS GEGDLSSI TDKKNRVSIRRDSTGLPLWLMGPDGQTFWFTGMTNNALKSVAACQGEIA VMTYHGSSGLLATKSNEDGWSTFYEDNYGRLTNVITYPTGRVSSYRTSDSTVRVQTE GSNKEDITVTTNLSASGTFYTLMQDOVKNSYIIGLDGSLRLVLANGMEVSLHTEPHLL SGTVNPTISKRNVTLPIDNGLNLVEWRQREQAQGVTVYGRRLRVHNRNLLSMDFDR VTRTEKVYDDHRKFTLRHYDHAGRP TLWAPSSRLNGVNTYSPGGHIAGIQRTMSV RMEYDQNGRITSKIFADGKSWSYTYLEKSMVLLLYSQRYIFEFDKNDRLSSVTMPNV ARQTLETTRSIGYRNTYRPEGNATVLQDISEDGLLQTIHQGTGRRVIYKYGKLSR LLEILYDTTTRIAFSYDESAGMLKTVGLQSEGFAC TIRYQIGPLIDRQIRFPSEGMV NARFDYNDNSFRVTSMQAVINETPLDILYRDVDSGTEQFGKFGVIYDINQIIT TAVMTHTKHFDAYGRVKEVQYEIFRSLMYVMVQFDNMGKRVAKELKVGPYANTRYA YEYDADGQLQVVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRDIRDIRLGDV QYRLDEDEGLRQGNDFEYNSAGLLVKYTNKVGWTKIYRYDGLGRRVSSRSTQGHH LQFFYADLSSPTRVTHMYNHSSEITSLYDYLQGHLFAMELSSGDFYVACDNIGTPL AVFSGAGLMKQILHTAFGEVYLDNSPFLVIGYQGGLYEPLTKLVHMGRRDYDLA GRWTTDPDHIRKRLNSDNI VPFNLYMFKNNNPLNSQYTKCYMTDVNSWLVTFGFLY NVIPGYRKPVTDAMEPSYELVHTQIKQEWDSKTSVLGVQCEVQQLKSFVRLERFGQ IYSASDSGCPPTPLHTLFATGTSLFGKGVKVAIREGRVEADIIISLANEDGRRIAA VLD KASYLQDLHFTIAGLDTHYFKVSGLVEGDLSTVGTQRTLETGTQVNTVTSQVNMVLG GRSRRITDIQMXYGTLNLNRYGSSVDEEKVRVLELARQAVATAWAHERHRLRQEE GSAWTDGERQQLLSGRVQGYEGFYIVSVDQFPELTDNINNHNHFWRQTEMGR"
ORIGIN	
Alignment	Scores:

Pred. NO.: 5.47e-68 Length: 9264
Score: 1495.50 Matches: 287
Percent Similarity: 70.34% Conservative: 45
Best Local Similarity: 60.81% Mismatches: 67
Query Match: 69.79% Indels: 73
DB: 5 Gaps: 3

US-10-029-020-14_COPY_1_400 (1-400) x AB026980 (1-9264)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAlaGluArg 20
Db 323 ATGGAAGTGAAGAACGACGACCCCTACCGCTCTCTGACCTCCAGCGGACACGGAGCGC 382
QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 383 CGCTACACCACTCATCCGCCGACAGCGAGGATGGCAAGATCAACCCCTAAATCTTACAGC 442
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 443 TCGAGTGAGACCCCTCAAAGCCCTTCGACCAAGGACTCCAGACTCGGCTTATGGCAGCCGCGTC 502
QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 503 AAAGACCTGGTGCACCATGAGGCGGACGAGTTCAGCAGACAAAGGCGCGGACTTTTCTCTC 562
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 563 AGAGACATGGCCCTTCGGAGATCCCGTCCGCCACATATGGGAGCGTACCGGACAGAAATG 622
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 623 GGCCCTCCCAACCGGACTACTCAGTGAGTGTGGCATCAGACGCTGACACAGAAACAGAC 682
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSer---ThrArgSerGly 139
Db 683 GGCATAATGTCCCCGAGACGCGCTCGCCTCTGGGGCCGACCAACCAATCCCGC 742
QY 140 ArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHis 159
Db 743 CGCAGTTCCCTGCTGTTTTCAGCAGAGCCAACTCCAACCTTACGCTCACCGACACCGAGCAT 802
QY 160 GluAsnThrGluThrAspHisPro----- 167
Db 803 GAAACACTGAGAATGTCCTCTCTGCAATGCTCATCTGCTCCTCTCTCTCTCTCTCTCTCTCT 862
QY 167 ----- 167
Db 863 TCCCTTTACCTTCCCTTCCCTTCCCAACGAGCAATCAGAGCCAGGGAAGTTGCTAGGTAAC 922
QY 167 ----- 167
Db 923 AGCGGGGCTCAGCGCGCGCGGACTCTGAGTCGGAAGATGAGTTTGGCCCAATTCATTTC 982
QY 168 -----Gly 168
Db 983 TTAGTTAAACCGGCTCAGGGAACGTCGTGCACTCCGCGCGCGCTACTGCTAATGAGGGC 1042
QY 169 GlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeuSerHisAlaHisThr 188
Db 1043 TCGTTCCAGAACTACTCAGGTTGCGAACGCCCACTTACCCCTGTTTTCATCTCGCACTCC 1102
QY 189 ProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArg 208
Db 1103 CCCAGTCAGCACCAACACCGCCTCCATCGGTTCCCTGAGCCGACGCAACTACACCCAGCGC 1162
QY 209 SerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGly 228
Db 1163 AGTAACCGGAGCCCGGACCCACCGATAGTTTACGCTCCCAATGAGGGTCCGACA---AGT 1219
QY 229 AlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGlu 248
Db 1220 GCGGAGGACTCCAGAGCGGCCAGGACAACTGGCTTCTCAACAGCAACGTTCTCTCTAGAG 1279

QY 249 ThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMet 268
Db 1280 ACCAGAAACATAGCAAGACAGACATTCTTAGAGACTTTGCAGGACAACTTTATAGAGATG 1339
QY 269 AspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys 288
Db 1340 GACATACTGGCGACAGCTCGCGCGACGGCGCTTACACTGATGGACACTTCTTTTCAAG 1399
QY 289 ProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 1400 CCTGGCGGACCTCGCCTTTGTACTGCACCACTTACCTGCTGCTTACCTTTGACGTCACG 1459
QY 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 1460 ACGGTGACTCTCTCTCCGCCCGCGCCCTGCCAAGAAACACCTTCTCACGGCCTGCCTTT 1519
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1520 AGCCTAAAGAAACCCCTATAAGCACTGCAACTGGAAATGTGCTGCCCTCAGTGCCATCCTC 1579
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1580 ATCTCCGTCACTCTGGTGTCTCTGTGGCATACTTTATTGCCATGCACCTCTTTGGACTC 1639
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSer 388
Db 1640 AACTGGCATTTGCAGCCGTCGAGGCGGAGGAGGAGATATATCAGTGCAGGAGGACAACACCA 1699
QY 389 SerTrpProValProThrAspValSerLeuTyrPro 400
Db 1700 GGCCTACATCTGCCACAGATCTGGGGCTGCCACCT 1735

RESULT 12
AR270936
LOCUS AR270936 2387 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 11 from patent US 6500941.
ACCESSION AR270936
VERSION AR270936.1 GI:29702178
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2387)
AUTHORS Schaefer,G.M. and Sliwowski,M.
TITLE Gamma-hergulin
JOURNAL Patent: US 6500941-A 11 31-DEC-2002;
FEATURES Location/Qualifiers
source
1. .2387
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 9.1e-58 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 59.96% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x AR270936 (1-2387)

QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCTGCGAGAACCAACCGCGGCTCCGACGCCCGCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACACCCCAACAGACCAACCGCGGCTCCATTAACTCCCTGAAACCGGGGCAACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCAGCCCGGCCCCCAGGACCACTCGCTCTCCGAGAGAGCCCCCT 180

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 170 TATAGTTCAGTGAAACCTTGAAAGCTTTCGATCATGATTATTACGGCTGCTTTATGGA 229

QY 58 SerArgValLysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsn 77
Db 230 AACAGAGTAAAGGATTGGTCCACAGAGAAGCCGACGAGTATACTAGACAAGCAGAGAT 289

QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArg 97
Db 290 TTTACCTAAGGCAGTTAGGAGTGTGTAATCCGCAACTCGAAGAGGAGTGGCATTTCTGT 349

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 350 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCCATCAGTGCAGGTCAGATCGGATACG 409

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 410 GAAACCAAGCAGTGTATCCCTGAGCATGCCATGAGACTTTGGGGCAGGGGTCAAA 469

QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 470 TCGGGCGCAGTTCTCGCTGTCAAGCCGCTCAACTCCGCCCTCACCTGACAGACACG 529

QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 530 GAGCACGAGAACAGGTGCGACAGTGAGAGCGAGCAACCTTCAAAACAACCCAGGGCAACC 589

QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 590 ACCCTGACGCCTTTGCCGCCA-----TCCCAACAGCAGCACCCCGCGCAGCATCAC--- 640

QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 641 CCGTCCATCACTTCCCTCAATAGAAACTCCCTGACCAATAGAGGAACACAGATCCGGCC 700

QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 701 CCG-----CCGGCTGCTTTGCCCGCGCAGCTGCAAAACCACA 736

QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 737 CCGGAGTCGCTCCAGCTGCAGGACAGCTGGGTCTCTGGCAGTAATGTACCCTGGAAAGC 796

QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 797 AGG----- 799

QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 800 -----CATTCCTATTCAAAACA 817

QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 818 GGGACAGGACGACGCCACTGTTCAGTACGGCAACCCCGGATACACAATGGCATCTGGC 877

QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 878 TCTGTTATTCTCCGCCTACCCGGCCACTTCCTAGAAACACCCCTATCAAGAAGTGCTTTT 937

QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 938 AAATTCAGAAGTCTTCAAAGTACTGCAGCTGGAGGTGCACCCGCACTGTGTGCTGATGGG 997

QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 998 GTCTCAGTGCTCCTGGCCATTCTCCTCTCCTATTTATAGCAATGCATCTATTGGCCTC 1057

QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1058 AACTGGCACTTACAGCAGACGCGAAATGACACATTCGAGAATGGAAGTGAATTCTGAC 1117

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1118 ACC-----GTGCCAACAAACACTGTATCGTTACCT 1147

RESULT 14
AF195420 1476 bp mRNA linear PRI 27-JAN-2000
LOCUS Homo sapiens ODZ3 (ODZ3) mRNA, partial cds.
DEFINITION AF195420
ACCESSION AF195420.1 GI:6760372
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1476)
AUTHORS Ben-Zur,T., Feige,E., Motro,B. and Wides,R.
TITLE The mammalian Odz gene family: Homologs of a Drosophila pair rule gene with expression implying distinct yet overlapping developmental roles
JOURNAL Dev. Biol. 117, 107-120 (2000)
REFERENCE 2 (bases 1 to 1476)
AUTHORS Ben-Zur,T., Motro,B. and Wides,R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel
FEATURES
source
1..1476
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..>1476
/gene="ODZ3"
/note="Ten-m3"
151..>1476
/gene="ODZ3"
/note="similar to the Drosophila melanogaster pair rule odd Oz/Ten-m gene sequence"
/codon_start=1
/product="ODZ3"
/protein_id="AAF28318.1"
/db_xref="GI:6760373"
/translation="MDVKERRPYCSLTKSRRKERRYNTSSADNEECRVPTQKSYSS
ETLKAFDHDSSRLLYGNRVKDLVHREADEFTROQNFTLRQLGVCEPATRRGLAFCAE
MGLPHRGYSISAGSDADTENEAVMSPEHAMRLWGRGVKSRSSCLSSRSNSALTLDPT
EHENKSDSENEQPSNQOSTLQPLPPSHKQHSQHPHSITSLNRNSLTNRNQSPAP
PAALPAELQTTPTESVQLQDSWVLGSNVPLESRHFLEPKTGTGTTPLFSTATPGYTMASG
SVYSPTRPLRPNTLSAFKFKKSKYCSWKCTALCAVGVSLLAILLSYFIAMHLF
GLNWQLQQTENDTFENGKVNDSMTPTNTVSLPSGDNKLGFTQENNTIDSGELDIGR
RAIQEIPPGIFWRSQLFIDQPFLKFNISLQKDALIGVYGRKKLPPSHTQ"

ORIGIN

Alignment Scores:
Pred. No.: 1.34e-42 Length: 1476
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 9 Gaps: 12

US-10-029-020-14_COPY_1_400 (1-400) x AF195420 (1-1476)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAAGAACGCGAGGCCTTACTGCTCCCTGACCAAGACAGACGAGAGGAA 210

QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGCGCTACACAAATTCTCCGACAGCAATGAGGAGTGCCGGGTACCCACAGAGATCC 270

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTCCAGCGAGACATTGAAAGCTTTTGATCATGATTCTCGCGCTGCTTTACGGC 330

QY 58 SerArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGGATTGGTTTCACAGAGAAGCAGACGAGTTCACTAGACAAGGACAGAAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTTAAGGCAGTTAGGAGTTTGTGAACCAAGCAACTCGAAGAGGACTGGCATTTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTGATGTCCCCAGAGCATGCCATGAGACTTTTGGGCAGGGGGGTCAAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCCGAGCTCCTGCCTGTCAAGTCGGTCCAACCTCAGCCCTCACCTGCACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAAAACAAGTCCGACAGTGAGAAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690
QY 175 ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGCCGCT-----TCCCATAGCAGCACTCTGCACAGCATCAT--- 741
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCCATCACTTCTCTCAACAGAAACTCCCTGACCAATAGAAGGAACCAAGATCCGGCC 801
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CGGGTGCTTTTGCCCGCGAGCTGCAAAACCACA 837
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCCGAGTCGGTCCAGCTGCAGGACAGCTGGGTCTTGGCAGTAATGTACCACTGGAAAGC 897
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTATTCAAAACA 918
QY 290 Gly----GlyThrSerProLeuPheCysThrThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAACGCCACTGTTTCAGTACTGCAACCCCGAGATACACAATGGCATCTGGC 978
QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTATTCAACCACTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTTT 1038
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAAGAAGTCTTCAAAGTACTGTAGCTGGAATGCACCTGCACCTGTGTGCCGTAGGG 1098
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCCTGGCAATACTCTCTTATTTATAGCAATGCATCTCTTTGGCCCTC 1158
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAAATGACACATTTGAGAATGGAAGAAGTGAATTCTGAT 1218
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAACACACTGTGTCAATTACCT 1248

RESULT 15

AX662355 8645 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 37 from Patent WO02062999.
DEFINITION AX662355
ACCESSION AX662355
VERSION AX662355.1 GI:29163217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millet,I. and Macdougall,J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 37 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source Location/Qualifiers
1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 9.6e-42 Length: 8645
Score: 987.00 Matches: 212
Percent Similarity: 62.89% Conservative: 49
Best Local Similarity: 51.08% Mismatches: 90
Query Match: 46.06% Indels: 64
DB: 6 Gaps: 12
US-10-029-020-14_COPY_1_400 (1-400) x AX662355 (1-8645)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAAGAACGCGAGCCTTACTGCTCCCTGACCAAGAGCAGACGAGAGGAA 210
QY 20 ArgArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGCGCTTACACAAATTCTCTCCGACACAATGAGGAGTGCCTGAGGAGTACCCACAGAGTCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCACGCGAGACATTGAAAGCTTTTGATCATGATTCCTCGGGCTGCTTACGGC 330
QY 58 SerArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn 77
Db 331 AACAGAGTGAAGGATTTGGTTTCACAGAGAAGCAGACGAGATTCACTAGACAAGGACAGAAT 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTTAAGGCAGTTAGGAGTTTGTGAACCAAGCAACTCGAAGAGGACTGGCATTTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTGATGTCCCCAGAGCATGCCATGAGACTTTGGGCAGGGGGGTCAAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCCGAGCTCCTGCCTGTCAAGTCGGTCCAACTCAGCCCTCACCTGCACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAAAACAAGTCCGACAGTGAGAAATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690

```
QY 175 ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db |||::: ||||| ||||| ||||| ||||| |||||
691 ACCCTGCAGCCCTTGCCGCT-----TCCCATAGCAGCACTCTGCACAGCATCAT--- 741

QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
742 CCATCCATCACTTCTCTCAACAGAAACTCCCTGACCAATAGAGGAAACCAGAGTCCGGCC 801

QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
802 CCG-----CGGCTGCTTTGCCCGCCGAGCTGCAAAACCACA 837

QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db ||| |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
838 CCCGATCCGTCAGCTGCAGGACAGTGGGTCTTGGCAGTAATGTACCACCTGGAAAGC 897

QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db ||| |||
898 AGG----- 900

QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
901 -----CATTCCTATTCAAAACA 918

QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db ||| |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
919 GGAACAGGTACAACGCCACTGTTTCAGTACTGCAACCCCGAGATACAAATGGCATCTGGC 978

QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db ::||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
979 TCTGTTATTACCCACTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTTT 1038

QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1039 AAATTCAAGAAAGTCTTCAAAGTACTGTAGCTGGAAATGCACTGCACTGTGTGCCGTAGGG 1098

QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db ::||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1099 GTCTCGGTGCTCCTGGCAATACTCCTGTCTTATTATTATAGCAATGCATCTCTTTGGCCTC 1158

QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1159 AACTGGCAGCTACAGCAGACTGAAAATGACACATTTGAGAATGGAAGTGAATTCTGAT 1218

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db ||| |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
1219 ACC-----ATGCCAACAAACACTGTGTCAATTACCT 1248
```

Search completed: August 14, 2004, 11:51:10
Job time : 7138.72 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 12:05:42 ; Search time 5140 Seconds
(without alignments)
3372.999 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76
Perfect score: 2132
Sequence: 1 MDVKERKPYRSLTRRRDAER.....EITEDTASSWPVPTDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFT0.spool/MITRA020/runat_06082004_114100_206/app_query.fasta.1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MITRA020 @CGN_1_1_3731@runat_06082004_114100_206 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2127	99.8	8354	6	AX556500	Sequence
2	2122	99.5	1680	6	AR270935	Sequence
3	2122	99.5	3111	6	BD270887	Method of
4	2122	99.5	3111	6	AR270934	Sequence
5	2122	99.5	3111	6	BD140153	Utilizati
6	2122	99.5	3111	9	AF009227	Homo sapi
7	2082	97.7	8438	6	AX675551	Sequence
8	2066	96.9	8645	6	AX600210	Sequence
9	2038	95.6	8585	10	AB025413	Mus muscu
10	1959	91.9	9722	10	AF059485	Mus muscu
11	1482.5	69.5	9264	5	AB026980	Danio rer
12	1285	60.3	2387	6	AR270936	Sequence
13	982	46.1	8964	10	AB025412	Mus muscu
14	975	45.7	8645	6	AX662355	Sequence
15	975	45.7	8675	6	AX662353	Sequence
16	975	45.7	8675	6	AX662353	Sequence
17	905	42.4	2496	5	GGA245711	Gallus ga
18	905	42.4	2496	6	AX250064	Sequence
19	904	42.4	8689	6	AX250067	Sequence
20	904	42.4	8689	10	AF086607	Rattus no
21	904	42.4	8797	6	AX250063	Sequence
22	904	42.4	8797	6	AX250066	Sequence
23	904	42.4	8797	10	AB025411	Mus muscu
24	885	41.5	8409	5	GGA279031	Gallus ga
25	885	41.5	8409	6	AX250068	Sequence
26	859	40.3	8575	6	AX921803	Sequence
27	828	38.8	9729	6	AX250013	Sequence
28	828	38.8	9826	6	AX250008	Sequence
29	777	36.4	8118	5	GGA238613	Gallus ga
30	734	34.4	8373	10	AB025410	Mus muscu
31	733.5	34.4	8297	9	AF100772	Homo sapi
32	641.5	30.1	8473	6	AX662357	Sequence
33	641.5	30.1	8487	6	AX662359	Sequence
34	522	24.5	1755	6	AX050264	Sequence
35	508	23.8	1727	6	AX714026	Sequence
36	508	23.8	1727	9	AK056053	Homo sapi
37	506.5	23.8	7713	9	HSM808325	Homo sapi
38	487.5	22.9	157677	2	AP002470	Homo sapi
39	487.5	22.9	166063	9	AP002515	Homo sapi
40	487.5	22.9	168308	2	AC084775	Homo sapi
41	474	22.2	87834	9	AP002957	Homo sapi
42	474	22.2	159946	2	AC079194	Homo sapi
43	458.5	21.5	93474	2	AC136801	Rattus no
44	458.5	21.5	173756	2	AC134948	Rattus no
45	458.5	21.5	245722	2	AC120288	Rattus no

ALIGNMENTS

```
AX556500
LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,
Smithson,G., Li,L. and Ji,W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source
1. .8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6.7e-103 Length: 8354
Score: 2127.00 Matches: 397
Percent Similarity: 99.25% Conservative: 0
Best Local Similarity: 99.25% Mismatches: 3
Query Match: 99.77% Indels: 0
DB: 6 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x AX556500 (1-8354)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 35 ATGGAGCTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCCGACCGCGAGCGC 94
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 95 CGCTACACCAAGCTCGTCCGCGGACAGCGAGGGCAAGCCCCCGCAGAAATCGTACAGC 154
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 155 TCCAGCGAGACCTGAAGGCCTACGACCGAGCGCCGCTAGCCTATGGCAGCCGCGTC 214
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 215 AAGGACATTGTGCCGCGAGGCGCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCTG 274
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 275 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCAGCGGACCCTGTACCGGACAGACATT 334
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 335 GGCCTGCCCAATGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 394
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 395 ACGGTGCTGTCCCTGAGCACCCCGTGGTCTGTGGGGCCGGAGCACACGCTCAGGGCGC 454
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 455 AGCTCCTGCCTGTCCAGCGGGGCCAAATTCCAATCTCACACTCACCGACACCGAGCATGAA 514
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 515 AACACTGAGACTGATCATCGGGCGGCCTGCAGAACACAGCGCGGCTCCGGACGCCCGC 574
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
```

```
Db 575 CCGCGGCTCTCGACGCCCAACACCCCCAACAGCACACCGGGCTCCATTAACTCCCTG 634
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 635 AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCACCGACCACTCGCTC 694
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 695 TCCGGAGAGCCCCCTGCCGCGCGGCCAGGAGCCTGCCACAGCCAGGAGAACTGGCTG 754
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 755 CTCACAGCAACATCCCCCTGGAGACCAGGAACCTTAGGCAAGCAGCCATTCCTAGGGACA 814
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 815 TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGCGCCTCCCGCCATGATGGGCTTAC 874
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 875 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCACATCA 934
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 935 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGCGC 994
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 995 AGCACCTTCGCCCGCGCGCCTTTAACTCAAGAAAGCCCTCCAAGTACTGTAACCTGGAAG 1054
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1055 TGCGCAGCCCTGAGCGCCATCGTCACTCTCAGCCACCTCTGGTCACTCTCTGTGGCATACTTT 1114
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1115 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1174
QY 381 GluIleThrGluAspThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400
Db 1175 GAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1234
RESULT 2
AR270935 AR270935 1680 bp DNA linear PAT 10-APR-2003
LOCUS Sequence 3 from patent US 6500941.
DEFINITION AR270935
ACCESSION AR270935
VERSION AR270935.1 GI:29702177
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1680)
TITLE Schaefer,G.M. and Sliwowski,M.
JOURNAL Gamma-hergulin
FEATURES Patent: US 6500941-A 3 31-DEC-2002;
Location/Qualifiers
1. .1680
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2.15e-103 Length: 1680
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 6 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x AR270935 (1-1680)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
```


QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	574	CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCTGTACCGACAGACATT	633
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	634	GGCCTCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATCCGACATGGAGGTGAC	693
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	694	ACGGTGCTGTCCCCCTGAGCACCCCGTGGCTGTGTGGGCGGAGACACGGTCAGGGCGC	753
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	754	AGCTCCTGCTGTCCAGCGGGCCAATTCCAATCTCACACTCACGACACCGAGCATGAA	813
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	814	AACACTGAGACTGATCATCCGGCGGCTGCAGAACCAACGCGGCTCCGGACGCGCGC	873
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu	200
Db	874	CCGCGCTCTCGACGCCACACCCCCAACAGACACACCGCGGCTCCATTAACTCCCTG	933
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	934	AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGCCCCACGGACCACTCGCTC	993
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	994	TCCGGAGAGCCCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1053
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
Db	1054	CTCAACAGCAACATCCCCCTGGAGACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1113
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	1114	TTGCAGGACAACTCATTTGAGATGGACATTCTCGGCGCTCCCGGCGGATGGGCTTAC	1173
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	1174	AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTTCACCACTCA	1233
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg	320
Db	1234	CCAGGTACCCACTGACGTCCAGCACAGTGACTCTCTCCCGCCCGACCCCTGCCCCG	1293
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys	340
Db	1294	AGCACCTTCGCGCGCGCGGCTTTAACCTCAAGAGCCCTCCAAGTACTGTAACTGGAAG	1353
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe	360
Db	1354	TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTTCATCTCTGTCATAC	1413
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1414	GTGGCCATGACACCTGTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGAGATGAT	1473
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	1474	GAGATCAGGAGGACACAGCCAGCAGTTGGGCTGTGCCAACCGACGCTCTCCCTATACCC	1533
RESULT 4			
AR270934			
LOCUS	AR270934	3111 bp	DNA linear PAT 10-APR-2003
DEFINITION	Sequence 1 from patent US 6500941.		
ACCESSION	AR270934		
VERSION	AR270934.1 GI:29702176		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unknown.		

Unclassified.

REFERENCE
1 (bases 1 to 3111)
AUTHORS
Schaefer,G.M. and Sliwkowski,M.
TITLE
Gamma-herregulin
JOURNAL
Patent: US 6500941-A 1 31-DEC-2002;
FEATURES
Location/Qualifiers
1..3111
source
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 4.2e-103 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 6 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AR270934 (1-3111)

QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	334	ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCGGAGCGC	393
QY	21	ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	394	CGCTACACACAGCTCGTCCGCGGACAGCGAGAGGGCAAGCCCCGAGAAATCGTACAGC	453
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	454	TCCAGCGAGACCCCTGAAGGCCTACGACACGAGACGCCCGCTAGCCTATGGCAGCGCGTC	513
QY	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80
Db	514	AAGGACATTGTCCCGCAGGAGGCGCGGAGGAATTCTGCCGACAGGTGCCAACTTCACCCCTG	573
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	574	CGGGAGCTGGGGCTGGAAGAAGTAACGCCCTTACCGGACCCCTGTACCGGACAGACATT	633
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	634	GGCCTCCCCACTGCGGCTACTCCAATGGGGCTGGCTCTGATGCCGACATGGAGGTGAC	693
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	694	ACGGTGCTGTCCCCCTGAGCACCCCGTGGCTCTGTGGGGCCGAGACACGGTCAGGGCGC	753
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	754	AGCTCCTGCTGTCCAGCGGGCCAATTCCAATCTCACACTCACGACACCGAGCATGAA	813
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	814	AACACTGAGACTGATCATCCGGGCGGCTGCAGAACCCACGCGGCTCCGGACGCCCGCG	873
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200
Db	874	CCGCGCTCTCGACGCCACACCCCCAACACAGCACCCACGCGGCTCCATTAACTCCCTG	933
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	934	AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGCCCCACGGACCACTCGCTC	993
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	994	TCCGGAGAGCCCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1053
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
Db	1054	CTCAACAGCAACATCCCCCTGGAGACCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
|||||
Db 1114 TTGCAGGACAACTCATTTGAGATGGACATTTCTGGCGCTCCCGCCATGATGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
|||||
Db 1174 AGTACGGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCACATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
|||||
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTAATCTCTCCGCCCGACCCCTGCCCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
|||||
Db 1294 AGCACCTTCGGCCCGCGGCTTTAACTCAAGAAGCCCTCCAAGTACTGTAAGTGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
|||||
Db 1354 TGGCAGCCCTGAGCGCCATCGTCACTCTCAGCCACTCTGGTCACTCTGTCATCTTT 1413

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
|||||
Db 1414 GTGGCCATGCACCTGTTGGCCTAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
|||||
Db 1474 GAGATCAGGAGGACACAGCCAGTGGCTGTGCCAACCGACGTCTCCCTATACCC 1533

RESULT 5

BD140153 3111 bp DNA linear PAT 18-SEP-2002
LOCUS Utilization of heregulin as epithelial cell growth factor.
DEFINITION BD140153
ACCESSION
VERSION BD140153.1 GI:23235098
KEYWORDS JP 2002509076-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 3111)
Sliwowski,M. and Kern,J.A.
Utilization of heregulin as epithelial cell growth factor
Patent: JP 2002509076-A 6 26-MAR-2002;
GENENTECH INC, UNIVERSITY OF IOWA RESEARCH FOUNDATION

COMMENT
OS Homo sapiens (human)
PN JP 2002509076-A/6
PD 26-MAR-2002
PF 03-FEB-1999 JP 2000530226
PR 04-FEB-1998 US 09/020598
PI MARK SLIWKOWSKI,JEFFREY A KERN
PC A61K38/00,A61K35/12,A61K39/395,A61K39/395,A61P11/00,A61P43/00,
PC C07K14/485,
PC C07K16/22,C12N15/09,C12P21/08,A61K37/02,C12N15/00 CC
Utilization of heregulin as epithelial cell growth factor FH Key
Location/Qualifiers
FT source 1..3111
FT /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..3111
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 4.2e-103 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 6 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BD140153 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
|||||
Db 334 ATGCACGTGAAGAGAGGAAGCCTTACCGCTCGTGTACCCGGCGCGACCGGAGCGC 393

QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
|||||
Db 394 CGCTACACCACTCGTCCCGGACACGAGGAGGGCAAGCCCCCGAGAAATCGTACAGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||||
Db 454 TCCAGCGAGACCTGAAGGCCTACGACAGGACGCCCGCTAGCCTATGGCAGCCGCGTC 513

QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
|||||
Db 514 AAGGACATTGTCCCGCAGGAGGCGGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCTG 573

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
|||||
Db 574 CGGAGCTGGGCTGGAAGAAGTAACGCCCTCACGGGACCTGTACCGGACAGACATT 633

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
|||||
Db 634 GGCCTCCCCCACTGGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
|||||
Db 694 ACGTGCTGTCCCTGAGCACCCCGTGGCTCTGTGGGGCCGAGACACCGTTCAGGGCGC 753

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
|||||
Db 754 AGTCTCGCTGTCCAGCCGGGCCAATTCCAATCTCACACTACCGACACCGAGCATGAA 813

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
|||||
Db 814 AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACCCACGCGGCTCCATTAACTCCCTG 873

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
|||||
Db 874 CCGCGGCTCTGCACGCCCAACACCCCAACCAGCACACCGCGGCTCCATTAACTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
|||||
Db 934 AACCGGGCAACTTTCACGCCGAGGAGCAACCCAGCCCGGCCCCCAGGACCACTCGCTC 993

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
|||||
Db 994 TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCCGAGGAGAACTGGCTG 1053

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
|||||
Db 1054 CTCACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGACA 1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
|||||
Db 1114 TTGCAGGACAACTCATTTGAGATGGACATTTCTCGGCGCTCCCGCCATGATGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
|||||
Db 1174 AGTACGGGCACTTCCTCTTCAAGCTGGAGGCACCTCCCGCTCTTCTGCACCACATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
|||||
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
|||||
Db 1294 AGCACCTTCGGCCCGCGGCTTTAACTCAAGAAGCCCTCCAAGTACTGTAAGTGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
|||||
Db 1354 TGGCAGCCCTGAGCGCCATCGTCACTCTCAGCCACTCTGGTCACTCTGTCATCTTT 1413

```
QY      361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
      |||
      1414 GTGGCCATGCACCTGTTGGCTAAACTGGCACCTGCAGCCGATGAGGGGCAGATGTAT 1473

QY      381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
      |||
      1474 GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 6
AF009227 LOCUS      AF009227      3111 bp      mRNA      linear      PRI 17-SEP-1997
DEFINITION Homo sapiens gamma-hereregulin mRNA, complete cds.
ACCESSION AF009227
VERSION AF009227.1 GI:2406643
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3111)
AUTHORS Schaefer,G., Fitzpatrick,V.D. and Sliwkowski,M.X.
TITLE Gamma-hereregulin: a novel heregulin isoform that is an autocrine
growth factor for the human breast cancer cell line, MDA-MB-175
JOURNAL Oncogene 15 (12), 1385-1394 (1997)
MEDLINE 97472144
PUBMED 9333014
REFERENCE 2 (bases 1 to 3111)
AUTHORS Schaefer,G., Fitzpatrick,V.D. and Sliwkowski,M.X.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1997) Protein Chemistry, Genentech, Inc., 460
Point San Bruno Blvd, South San Francisco, CA 94080, USA
FEATURES
source
1..3111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22-p11"
/cell_line="MDA-MB-175"
/cell_type="breast cancer cell"
334..2640
/codon_start=1
/product="gamma-hereregulin"
/protein_id="AAC51756.1"
/db_xref="GI:2406644"
/translation="MDVKKRPYRSLRRRDAERRYTSSSADSEEGKAPQKYSSET
LKAYDQDARLAYGSRVKDIVQEAEEFCRTGANFTLRELGLSEVTPPHGTLRYTDIGL
PHCGYSMGAGSDADMEADTVLSPHVPRLWGRSTRSGRSSCLSSRANSNLTLDTEHE
NTETHPGGLQNHARLRTPPPLSHAHTPNQHHAAASINSLRGNFTPRSNPSPAPTDH
SLSGEPPAGGAQEPAAHQENWLNLSNIPLETNRNLGKQPLFLTQDNLIEMDILGASRH
DGAYS DGHF LFKPGGTSPLFCTSPGYPLTSSTVSPPPPLPRSTFARPAFNLKPS
KYCNWKAALSAIVISATLVLLAYFVAMHLFGLNHLQPMEGQMYEITEDTASSWPV
PTDVS LPSGGTGLET PDRKGKTGTEGKPSFFPEDSFDSGEIDVGRRASQKIPPGT
FWRSQVFIDHPVHLKFNVS LKGAALVGIYGRKGLPPTQFDVVELLDGRRLLTQEAR
SLEGT PRQSRGT VPPSSHETGFIQYLDSGIWHLAFYNDGKSEVVSFLT TAIALPRL
KEMKQESAAGSKLVLCETSSSEYSSLRFKWFKNGNELNRKNKPNQIKQKPGKSEL
RINKASLADSGEYMKCVISKLGNDSSANITIVESNEIITGMPASTEGAYVSSSEPIR
ISVSTEGANTSSSTSTGTGSHLVKCAEKEKTFVCVNGGECFMVKDLSNPSRYLCKCP
NEFTGDRQCQNYVMASFYSTSTPFLSLPE"
ORIGIN
Alignment Scores:
Pred. No.: 4.2e-103 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 9 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x AF009227 (1-3111)
QY      1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
      |||
      1414 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGAGGGGCAGATGTAT 1473
```

```
Db      334 ATGGACGTGAAGGAGAGAAAGCCTTACCGCTCGCTGACCCGGCGCGCGACCCGAGCGC 393
QY      21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
      |||
      394 CGCTACACCAAGCTCGTCCCGCGGACGCGGAGGGGCAAAGCCCGCAGAAATCGTACAGC 453
QY      41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
      |||
      454 TCCAGCGAGACCTGAAGGCCTACGACCGACGCGCCCGCTAGCTATGGCAGCGCGTC 513
QY      61 LysAspIle***proGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
      |||
      514 AAGGACATTGTGCGCAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCTG 573
QY      81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
      |||
      574 CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 633
QY      101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
      |||
      634 GGCCTCCCCCACTGCGGCTACTCCATGGGGGTGGCTCTGTATGCCACATGGAGGCTGAC 693
QY      121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
      |||
      694 ACGGTGCTGTCCCTGAGCACCCCGTGGTCTGTGGGCGCGAGCACACGGTCAGGGCGC 753
QY      141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
      |||
      754 AGTCTCTGCTGTCCAGCGCGGCCAATTCGAATCTCACACTCACCGACACCGAGCATGAA 813
QY      161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
      |||
      814 AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCAACGCGCGGCTCCGGACGCCCGC 873
QY      181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
      |||
      874 CCGCCGCTCTCGCACGCCCAACACCCCAACACGACACCGCGGCTCCATTAACTCCCTG 933
QY      201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
      |||
      934 AACCGGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCGACGAGCACTCGCTC 993
QY      221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
      |||
      994 TCCGGAGAGCCCCCTGCGCGCGCGGCCAGGAGCCTGCCACGCGCGAGGAGAACTGGCTG 1053
QY      241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
      |||
      1054 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCATTCTCTAGGACA 1113
QY      261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
      |||
      1114 TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGCGCCTCCCGCCATGATGGGCTTAC 1173
QY      281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
      |||
      1174 AGTGACGGGCACTTCTCTTCAAGCCTGGAGGCACCTCCCGGCTCTTCTGCACCATCA 1233
QY      301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
      |||
      1234 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 1293
QY      321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
      |||
      1294 AGCACCTTCGCCCGCGCGCCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1353
QY      341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
      |||
      1354 TGCGCAGCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTTCATCTCTGTCATACTTT 1413
QY      361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
      |||
      1414 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGAGGGGCAGATGTAT 1473
```


QY 381 GluIleThrGluAspThrAlaSerSerTyrProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACAGCCAGCAGTTGGCTGTGCCAACCGACGTCCTCCCTATACCCC 1533

RESULT 7
LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO02055704.
ACCESSION AX675551
VERSION AX675551.1 GI:293333552
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
Stone,D.J.

TITLE Proteins, polynucleotides encoding them and methods of using the
same

JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)

FEATURES
source Location/Qualifiers
1..8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.6e-100 Length: 8438
Score: 2082.00 Matches: 395
Percent Similarity: 98.01% Conservative: 0
Best Local Similarity: 98.01% Mismatches: 5
Query Match: 97.65% Indels: 3
DB: 6 Gaps: 1

SEQ14-X-AT-28-64-76 (1-400) x AX675551 (1-8438)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 4 ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCGGACGCCGAGCGC 63

QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 64 CGCTACACAGCTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCGCGCAGAAATCGTACAGC 123

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 124 TCCAGCGAGACCCCTGAAGGCTTACGACCAGGACGCCCGCTAGCCTATGGCAGCCCGCGTC 183

QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 184 AAGGACATTGTGCCGCGAGGAGCGGAGGAATCTGCCGACAGGTGCCAACTTCACCCGTG 243

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 244 CGGAGCTGGGGCTGGAAGAGTAAAGTAAACGCCCTCACGGGACCTGTACCGACAGACATT 303

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 304 GGCCTCCCCACTGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGSCTGAC 363

QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 364 ACGGTGCTGTCCCTGAGCACCCCGTGCCTCTGTGGGGCGGAGCACACGGTTCAGGGCGC 423

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 424 AGCTCCTGCTGTCCAGCCGGGCCAATTCCAAATCTCACACTCACCGACACCGAGCATGAA 483

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 484 AACACTGAGACTGATCATCCGGCGGCTGCAGAACACACGCGGCTCCGGACCGCGCG 543

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 544 CCGCCGCTCTCGCACGCCACACCCCCAACACAGCACCGGGCCTCCATTAACTCCCTG 603

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 604 AACCGGGGCACTTCAGCCGAGGAGCAACCCAGCCGCCGCCACGACCACTCGCTC 663

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTyrLeu 240
Db 664 TCCGGAGAGCCCTCGCGCGCGGCCAGAGCCTGCCACGCCAGGAGAACTGGCTG 723

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 724 CTCAACAGCAACATCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTAGGACA 783

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 784 TTGCAGGACAACTCATTGAGATGGACATCTCGGCGCCTCCCGCCATGATGGGCTTAC 843

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 844 AGTGACGGGCACCTCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCATCA 903

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 904 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 963

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTyrLys 340
Db 964 AGCACCTTCGCCTGGCCGCCCTTTAACTCAAGAAGCCCTCCAACTACTGTAAGTGAAG 1023

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1024 TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGTGTCTCTGTGGCATACTT 1083

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTyrHisLeuGlnProMetGluGlyGlnMet --- 379
Db 1084 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCCCATGGAGGGGCAGATGACG 1143

QY 380 ---TyrGluIleThrGluAspThrAlaSerSerTyrProValProThrAspValSerLe 398
Db 1144 GATTATGAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCT 1203

QY 398 uTyrPro 400
Db 1204 ATACCCC 1210

RESULT 8
LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Wallia,N.K.

TITLE Proteins associated with cell growth, differentiation, and death

JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)

FEATURES Location/Qualifiers

source 1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7488573CB1"

ORIGIN

Alignment Scores: 1.14e-99 Length: 8645
Pred. No.: 2066.00 Matches: 390
Score: 2066.00
Percent Similarity: 97.50% Conservatives: 0
Best Local Similarity: 97.50% Mismatches: 6
Query Match: 96.90% Indels: 4
DB: 6 Gaps: 2

SEQ14-X-AT-28-64-76 (1-400) x AX600210 (1-8645)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 117 ATGGACGTGAAGGAGAGGAAGCCTTACCGCTGCTGACCCCGGCCCGCAGCGCGC 176
QY 21 ArgTyrThrSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 177 CGCTACACCAAGTCGTCGCGGACAGCGAGGGGCAAGCCCCCGCAGAAATCGTACAGC 236
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 237 TCCAGCGAGACCGCTGAAGGCCTACGACCGACGCGCCCGCTAGCCTATGGCAGCGCGTC 296
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 297 AAGGACATTGTGCGCAGGAGGCGGAGGAATCTGCGCACAGGTGCCAACTTCACCGCTG 356
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 357 CGGAGCTGGGGCTGAAGAAGTAACGCCCCCTCACGGGACCCTGTACCGACAGACATT 416
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 417 GGCCTCCCCACTCGCGCTACTCCATGCGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 476
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 477 ACGGTGCTGTCCCTGAGCACCCCGTGCGTCTGTGGGCGCGGAGCACACGCTCAGGGCGC 536
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 537 AGCTCCTGCCTGTCCAGCGGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 596
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 597 AACACTGAGACT-----CCGGGCGGCTGCAGAACACGCGCGGCTCCGGACGCGCGCG 650
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 651 CCGCGCTCTCGCACGCGCCACACCCCAACCCAGCACCGCGGCTCCATTAACTCCCTG 710
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 711 AACCGGGCAACTTTCACGCGGAGGAGCAACCCAGCGCGGCCCGCCACGACCGACTCGCTC 770
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 771 TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCCGCAGGAGAACTGGCTG 830
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 831 CTCACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTCCTAGGACA 890
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280

Db 891 TTGCAGGACAACCTCATTTGAGATGGACATTTCTCGGCGCCTCCCGCATGATGGGCTTAC 950
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 951 AGTGACGGGCACTTCTCTTCAAGCCTGGAGGCACCTCCCGGCTCTTCTGCACCACATCA 1010
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 1011 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCTGCCCGC 1070
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1071 AGCACCTTCGCGCGCGCGCTTAACTCAAGAGGCCCTCCAAGTACTGTAACTGGAAG 1130
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1131 TCGCGAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCATCTCTGTCATCTGCTGGCATACTTT 1190
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1191 GTGGGTAAGCACCTCTTC-----AACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1244
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1245 GAGATCACGAGGACACAGCCAGCTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1304

RESULT 9

AB025413 8585 bp mRNA linear ROD 08-MAY-1999
LOCUS Mus musculus mRNA for Ten-m4, complete cds.
DEFINITION AB025413
ACCESSION AB025413
VERSION AB025413.1 GI:4760781
KEYWORDS Ten-m4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

AUTHORS Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T., Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues
JOURNAL J. Cell Biol. (1999) In press
REFERENCE 2 (bases 1 to 8585)
AUTHORS Oohashi,T.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho, Okayama, Okayama 700-8252, Japan
(E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128, Fax: +81-86-222-7768)

FEATURES

Location/Qualifiers
1..8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1..8585
/gene="ten-m4"
188..8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"

gene

CDS

/translation="MDVKERKPYRSLRRRDAERRYTTSSADSEEGKGPQKYSSET
LKAYDQDARLAYGSRVKDMVPQEAEEFCRTGNTFTLRELGLGEMTPPHGTLRTDIGL
PHCGYSMGASDADLEADTVLSPHPVRLWGRSTRSGRSSCLSRANSNLTLDTEHE
NTETDHPSSLQNHPRILRTPPPLPHAHTPNQHHAAASINSLNRGNFTPRSNPSAPT
SLSGEPPAGSAQEPHTAQDNWLLNSNIPLFTRNLGKQPFGLTQDNLIEMDILSASRH

DGAYS DGHFLFKPGGTSPLFCTTSPGYPLTSTSVYSPPPRLPRSTFSRPAFNLKKPS
KYCNWKCAALSAILISATLVILLAYFVAMHFLGLNWNHLQPMEGQMOMEYITTEDTASSW
PVPTDVSLYPSGGTLETPDRKGAEGKPSLSLFPEDSFIDSGEIDVGRRASQKIPP
GTFWRSQVIDHPVHLKFNVSGLKAAALVGIYGRKGLPPSHTQFDFVELLDGRRLLTQE
ARSLGEPQQRQSRGPVPPSSHETGFIQYLDSGIWHLAFYNDGKESEVVSFLTTAIESVD
NCPNSCYNGNDCISGTCFLGFLGPDCGRASCPLVCSNGQYMKGRCLCHSGWKGAEE
CDVPTNQCIDVACSSHGTCIMGTCICNPGYKGESCEEVDCMDPTCSSRGVCRGECCH
SVGWGTCNETPRATCLDQCSGHTFLPDITGLNCDCPSWTGHDCSIETCAADCGGHGV
CVGCTCRCEDGWMGAACDQACHPRCAEHGTICRDGCKCESPCGWNGEHCTIAHYLDRVV
KEGCPGLCNGNGRCTLDLNGHVCVQLGWRGTGCTMETGCGDGLNDGDLVDCMD
PDCCLQPLCHVNPCLGLSPDPLDIIQETQAPVQQNLNPFYDRIKFLVGRDSTIPG
ENPFDDGGHACVIRGQVMTSDGTPLVGVNISFINNPLFGYTI SRQDGSFDLVTNGGISI
ILRFERAPFITQEHTLWLPWDRFFVMETIIVMRHEENETIPSCDLSNFAFPNFWVSPSL
TSFASSCAEKGPVPEIQALQEEIIVIAGCKMRLSYLSRTPGYKSVLRI SLTHPTIPF
NLMKVHLMVAVEGRILFRKWFAPADLSYFIWDKTDVYNQKVFGESEAFVSUVGYEYES
CPDLILWEKRTAVLQGYEIDASKLGGWSLDKHALNIQSGILHKNGENQFVSQQPPV
IGSIMNGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNIRRIFFPSGNV
TNI LEMNRKDFRHSHPAHKYYLATDPMGSAVFLSDTNSRRRVFKVKS TTVVVKDLVKN
EVVAGTGQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRVQNG
IIISTLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLVLDNNVVLLQISE
NHQVRIVAGRPMHCQVPGIDHFLLSKVATHATLESATALAVSHNGVLVIAETDEKKIN
RIRQVTTSGEISLVAGAPSGCDCKNDANCDFSGDDGYAKDAKLNTPSSSLAVCADGEL
YVADLGNIRIRFIRKNKPFNTQNMYELSSPIDQELYLFDTSKHLYTQSLPTGDYLY
NFTYTGDGDI THITDNNGMNVNRDSTGMLVLVDPDQGVYVWVTGNTSALRSVTTQ
GHELAMMTYHGNSGLLATKSNENGWTFYEYDSFGRLLTNVTFPTQVSSFRSDTSSV
HVQVETSSKDDVTIITNLSASGAFYTLLDQVQRNSYYIGADGSLRLLLLANGMEVALQT
EPHLLAGTVNPTVGKRNVTLPIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLS
LDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVTVSPGGHIAQIR
GIMSERMEYDQAGRITSIRIFADGKMWSYTYLEKSMVLHLHSQRQYIPEFDKNDRLSSV
TMPNVARQPLETIRSVGYARNIYQPPGNASVIQDFTEDGHLHHTFYLGTGRRVIYKY
GKLSKLAETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCTIRYRQIGPLIDRQIFRET
EEGMVNARFDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYYDI
NQIIITAVMTHTKHFDAYGRMKEQVEIFRSLMYMTVQYDNMGRVVKELXVGPYAN
TTRYSVEYDADGQLQTVSINDKPLWRYSYDLNGLNHLHLLSPGNSARLTPLRYDLDRIT
RLGDVQYKMDDEDGLFQRQGDVFEYNSAGLLIKAYNRASGWSVRYRYDGLRRVSSKS
SHSHLQFFFYADLTNPTKVTHLVNHSSEITSLYYDLOGLHFA MELSSGDEFYIACDN
IGTPLAVFSGTGLMIKQILYTAGYRIYMDTNPNFQIIIGVHGGLYDPLTLKLVHMGRRD
YDVLAGRWTPDHLEWKRLSSNIVPFHLYMFKNNNPISNQDIKCFMTD VNSWLLTF
GFQLNHVPYGPYKPDTDAMEPSVELVHTQMTKQEWDNKSILGVQCEVQKQKAFVTL
ERFDQLYGSTITSQQAPETKKEFASSGSIFGKGVKFALKDGRVTTDIIISVANEDGRII
AAILNNAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGRRTLENGVNVTVSQI
NTWLSGRTRRYTDIQLQYRALCLNTRYGTTVDEEKRVLELARQRAVRQAWAREQORL
REGEEGLRAWTGDGEKQQVLNTGRVQGYDGFVTSVEQYPELSDSANNIHFMRQSEMGR
R"

ORIGIN

Alignment Scores: 3.4e-98 Length: 8585
Pred. No.: 2038.00 Matches: 380
Score: 96.02% Conservative: 6
Percent Similarity: 94.53% Mismatches: 14
Best Local Similarity: 95.59% Indels: 2
Query Match: 10 Gaps: 1
DB:

SEQ14-X-AT-28-64-76 (1-400) x AB025413 (1-8585)

Qy	1	MetAspVallysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg	20
Db	188	ATGGACGTGAAGGAGAGAGAGCCCTATCGCTCGCTGACCCGGCGTCGCGATGCTGAGCGC	247
Qy	21	ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	248	CGCTACACAGACTCATCGGCAGACACGCGAGGAGGCAAGGGCCACAGAGTCCTACAGC	307
Qy	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	308	TCCAGTGAGACCCCTGAAGGCATATGACCAAGATGCTCGCCTAGCCTACGGCAGCCGCGTC	367
Qy	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80
Db	368	AAGGACATGGTACCACAGGAGGCCGAGAGTTCTGCGCGACAGGCACTAATTTCACTCTT	427
Qy	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100

Db	428	CGTGAAGTGGGACTGGGAGAGATGACGCCCCCTCATGGGACTCTCTACAGGACAGACATC	487
Qy	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	488	GGCCTCCACACTGTGGCTATTCCATGGGGGCCAGCTCTGATGCAGACTTGGAAGCAGAC	547
Qy	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	548	ACTGTGCTATCCCTGAACACCCAGTGCCTGTGGGCGGAGCACACGGTCAGGGCGC	607
Qy	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	608	AGTTCTGCCTGTCCAGCGGGCCAACTCCAACCTCACACTCAGGACACAGAGCATGAG	667
Qy	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	668	AACACAGAAACCGATCATCCAGCAGCCCTGCAGAACCAACCCCTCGGCTCCGGACGCCTCCT	727
Qy	181	ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu	200
Db	728	CCACCACTGCCCCACGCCCATACCCCAACCAAGAGCCAGCCCTCCATCAACTCCTCTTG	787
Qy	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	788	AACAGGGGCAACTTACCCCAAGGAGCAACCCAGCCCAACCCATGCCAGGACAACCTGGCTA	847
Qy	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	848	TCTGGGAGCCCCAGCAGGCGAGCGCCAGGAGCCAAACCCATGCCAGGACAACCTGGCTA	907
Qy	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
Db	908	CTCAACAGTAACATCCACTTGAAACCCAGAAACCTAGGCAAGCAGCCATCTCTAGGGACA	967
Qy	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	968	TTGCAGGACAACTCATTTGAGATGGACATTTCTCAGCGCCTCCCGCCATGATGGGCTTAC	1027
Qy	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	1028	AGTGACGGGCACTTCTCTTCAAGCCCGGAGGCACCTCCCCCGCTCTTCTGCACCATCC	1087
Qy	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320
Db	1088	CCAGGGTACCCCTAAAGTCTAGCACCGTGTACTCGCCCCACCCCGGCCCTTGCCCCGC	1147
Qy	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340
Db	1148	AGCACCTTCTCCCGACCAGCCCTTTAACTCAAGAAGCCTTCCAAGTACTGCAACTGGAAG	1207
Qy	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1208	TGTGACGCGCTGAGCGCCATCTCTCATCTCAGCTACGCTCGTCTCATCTCTCGCATACTTT	1267
Qy	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGly-----Gln	378
Db	1268	GTGGCCATGCACCTGTTTGGCCTAAAACTGGCACCTGCAGCCGATGGAGGGCAGATGCAG	1327
Qy	379	MetTyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeu	398
Db	1328	ATGTATGAGATCACGGAGGACACAGCCAGCAGCTGGCCCTGTGCCAACGGACGCTCTCCCTG	1387
Qy	399	TyrPro 400	
Db	1388	TATCCC 1393	

RESULT 10
AF059485
LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998
DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614

Db	1243	TTAGTTAAAAAGTGGCTCCGCGCAGCCTGGGGTCCGAGCAACGATCATCCCGACGACCTG	1302
QY	171	GlnAsnHisAlaArgLeuArgThrProProProLeuSerHisAlaHisThrProAsn	190
Db	1303	CAGAACCACTCGGCTCGGACGCCTCCTCCACCACTGCCCAACGCCCATACCCCAAC	1362
QY	191	GlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn	210
Db	1363	CAGCATCACGGCGCTCCATCAACTCCTTGAAACAGGGCAACTTCACCCCAAGAGCAAC	1422
QY	211	ProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln	230
Db	1423	CCCAGCCCAGCACCCACAGACCACTCACTCTCTGGGAGCCCCCAGCAGGCGCCACG	1482
QY	231	GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArg	250
Db	1483	GAGCCAACCCATGCCAGGACAACTGGGTACTCAACAGTAAATCCCACTGGAAACCCAGA	1542
QY	251	AsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIle	270
Db	1543	AACCTAGGCAAGCAGCCATTCTTAGGGACATGGCAGGACAACCTCATTTGAGATGGACATT	1602
QY	271	LeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysProGly	290
Db	1603	TTCAGCGCCTCCCGCGTGATGGGCTTACAGTGACGGGCACCTCTTYYTCAAGCCCGGA	1662
QY	291	GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal	310
Db	1663	GGCACCTCCCGCTCTTCTGCACCACATCCCCAGGGTACCCCTAAACGCTAGCACCGTG	1722
QY	311	TyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu	330
Db	1723	TATTCGCCCCCACCCCGCCCTGCCCCGAGCACTTTCTCCCGACCACTTTAAACCTC	1782
QY	331	LysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSer	350
Db	1783	AAGAAGCCTTCCAAGTACTGCAACTGGAAAGTGTGCAGCGCTGAGCGCCATCTCATCTCA	1842
QY	351	AlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrp	370
Db	1843	GCTACGCTCGTCATCCTGCTCGCATACTTTGTGGCCATGCACCTGTTGGCCTAAACTGG	1902
QY	371	HisLeuGlnProMetGluGly-----GlnMetTyrGluIleThrGluAspThrAlaSer	388
Db	1903	CACCTGCAGCCGATGGAGGGCAGATGCAGATGTATGAGATCACGGAGGACACAGCCAGC	1962
QY	389	SerTrpProValProThrAspValSerLeuTyrPro	400
Db	1963	AGCTGGCCTGTGCCAACGGACGCTCTCCCTGTATCCC	1998
RESULT 11			
AB026980			
LOCUS			
DEFINITION			
AB026980			
ACCESSION			
AB026980.1			
VERSION			
ten-m4.			
KEYWORDS			
SOURCE			
ORGANISM			
Danio rerio (zebrafish)			
Danio rerio			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;			
Cypriniformes; Cyprinidae; Danio.			
1 (sites)			
Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.			
AUTHORS			
TITLE			
Compartmentalized expression of zebrafish ten-m3 and ten-m4,			
homologues of the Drosophila ten(m)/odd Oz gene, in the central			
nervous system			
Mech. Dev. 87 (1-2), 223-227 (1999)			
JOURNAL			
MEDLINE			
99425191			
PUBMED			
10495292			
REFERENCE			
2 (bases 1 to 9264)			
AUTHORS			
Mieda,M.			
TITLE			
Direct Submission			

JOURNAL	Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute, RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hiroosawa, Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp, Tel:81-48-467-9713, Fax:81-48-467-9714)
COMMENT	Sequence updated (29-Jun-1999).
FEATURES	Location/Qualifiers
	1..9264
source	/organism="Danio rerio"
	/mol_type="mRNA"
gene	/db_xref="taxon:7955"
	/dev_stage="embryo"
CDS	1..9264
	/gene="ten-m4"
	323..8797
	/gene="ten-m4"
	/note="similar to Drosophila melanogaster tenm/odz and Mus musculus Doc4"
	/codon_start=1
	/product="ten-m4"
	/protein_id="BAA81893.1"
	/db_xref="GI:5307785"
	/translation="MEVKERRPYRSLTSRRDTERRYTSSSADSEDGKINPKSYSSET LKAFDQSRLAYGSRVKDLVHHEADSFSRQGPDFSLRDMAFGDPVPPHMGAYRTEMGL PHRDYSVVASDADTETDGMSPHAEVRLWGRSNTKSGRSSCLFSPRANSNLTLTDTTEH ENTENGPPLHCSASSSPVSPYPSPSHAANQSQGRLLNGSQAQAGRDSSEDEFQPN SFLVKTSGNVCTPAAATANECSFQNHSLRTPPLPLFHSKSPSQHHTASIGSLSRSN YTORSNPSPAPTSSAPNEGPTSAQDSSAQDNWLLNSNVPLETRNIAKQTFLETLQD NFIEDILATARRDGAYTDGHFLFKPGGTSPLYCTTSPGYPLTSSTVSPPPRPLPRN TFSRPAPFSLKKPYKHCNWKCAALSAILISVTLVFLAYFIAMHLFGLNWHLPQVQRQI YQLTEDNTSGLHLPDLDGLPLGNTGLEFPDRGSRDDGKLDGFFPEDSFIDMGEIDVG RKVAQLIPPGIFWRSQVFIDHPMYLKFNVLSKDALVGIYGRRLPPSHTQDFVELL DGRLLSQGLPGLDGPFPQAQORSLPVITSHDTGCIQYMDSGIWHLAVYNDGKETQOV SELTIAIDSIDDCPSNCFNGDCVSGNCHCFPGFRGPCDSCRASCPVLCSGNGQYLKGR CMCHSGWKGSEC DVPTNQCIDITCSGHGTCIVGTCICNPSPYKGENCEEVDCLDPTCSG RGVCVRGECHCFVGGWGPCESPRASCMEQCCHGSHGSLADTNTCNDHNWTGHDCSTE LCAADCGGHGICVAGSCRCDEGWMGTGCEQACHPRCSEHGTCCKGCECSPGWNGEH CTIEGCPGLCNGNRCITLGNNGWYCVQLGWRGAGCDTSMETACSDGKDNDDGLTDC MDPDCCLOASCHTTTSLCVGSPDPLDIQETQISSSLSTLQSFYQRIHFLVGRDSTHVI PDVNPFDGIHACVIRGQVVTSDGTPLVGNISFINFKPAYGTYITRQDGSFDLVNNGGV AIGLRFERAPFITQEHTLMLPWGRFFVMDTIVMRHEVNDIPSCDLSSTFRPMPVILPA PLTAFAGTCPERGIVVPEIQTLQEEVRIPGTMDRLXYLSRSTSGYKSLRLITLTHSTI PPSLMKVHLMVAVEGRLFRKWFSAAPNLSDYFVMDKTDVYSQKVGSLSEAFVSQFVEY ESCPDLILWEKRTAVLQGYETTASNLGGWSVDKHALNIQSGILHKGNGENIFISQOP PVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALACSDGSLYVGFNYYRRIFTTG NVTSVLELSNSPAHKYILATSPVSGWLYLSDTSSRKVFVKSLYAVKDVAKNLELVAG TGDQCLPYDETRCGDGGKAVEATLTPRGITVDKYGVIFVVDGTMIRRIDQNGIISTL LGFNDLTTSARPLSCDSVMDISQVRLEWPTDLAVSPMDNSLYVLDDNVVLQISENHQVR IVAGRPMHCQVPGLDHFLVSKIAIHATLESANALAVSHNGLLYIAESDEKKINVRQV STNGEISLLAGAPSGCDCKNDANCDCYSGDDGYAKDAKLNAPSSLAVSPDGELFIADL GNIRIYVRRNKAFNLNMYEISSPIDDELYLFDVNASHVFTQSLTTGDYLYNFTYS GEGDLSSITDKNKNRVSIRRDSTGLPLWLMGPDGQTFWFTGTNNALKSVAACQOGEIA VMTYHGSSGLLATKSNEDGWSTFYEYDNYGRLTNNVTYPTGRVSSYRTSDSTVRVQTE GSNKEDITVTTNLSASCTFYTLMDQVKNSYYIGLDSLRLLVLANGMEVSLHTEPHLL SGTVNPTISKRNVTLPIDNGLNLVEWRQKEAQGVTVYGRRLRVHNRNLLSMDFDR VTRTEKVYDDHRKFTLRHYDHAGRPTLWAPSSRLNGVNVITYSPGGHIAGIQRGTMVSV RMEYDQNGRITSKI FADGKSWSYTYLEKSMVLLLYSQRYIFEFDKNDRLSSVTMPNV ARQTLTETRSIGYRNTRYRPEGNATVLQDISEDGLLQTIHQGTGRRVIYKYGLSR LLEILYDTRIAFSDYDESAGMLKTVGLQSEGFACTIRYRQIGPLIDRQIFRSEEGMV NARFDYNDNSFRVTSMAVINETPLPIDLYRDYDVSGKTEQFGKFGVIYYDINQIIT TAVMTHTKHFDAVRVKEVQYEIFRSLMYMMVQFDNMGVRVAKELKVGPIYANTRYA YEYDADGQLQVVSINDKPLWRYSDNLNGLHLLSPGNSARLTPLRYDIRDITRLGDV QYRLDEGDGFLRQRGNDFFEYNSAGLLVKTYKNVNGWTIKYRYDGLGRRVSSSTQGGH LQFFYADLSSPTRVTHMYNHSSEITSLYDYLQGLHFAMELSSEDEFYACDNIQTPL AVFSGAGLMIKQILHTAFGEVYLDNSNPSFQLVYGGQGLYEPLTKLVHMGRRDYDVL GRWTTDPDHDIRKRLNSDNIVPFNLVFNKNNPLNSQOETKCYMTDVNSWLVPFGFLY NVIPGYRKPVTDAMEPSYELVHTQIKTQEWDSKSVLGVCQEVQRQLKSFVRLERFGQ IYSASDSGCPPTPLHTLFATGTSLFGKGVKVAIREGRVEADIIISLANEDGRRIAAVLD KASYLQDLHFTIAGLDTHYFVKSGLVGEGDLSLLGTVMGTQVGTLETGNVTVSQVNMVLG GRSRRITDIQMVGQITSLNVRYGSSVDEEKVRVLELARQAVATAWAHERHLRQGEE GSAWTDGERQQQLSSGRVQGYEGFYIVSVDPFELTDNINNHFWRQTEMGRR"
ORIGIN	
Alignment Scores:	

```
Pred. No.: 7.12e-69 Length: 9264
Score: 1482.50 Matches: 285
Percent Similarity: 69.92% Conservative: 45
Best Local Similarity: 60.38% Mismatches: 69
Query Match: 69.54% Indels: 73
DB: 5 Gaps: 3

SEQ14-X-AT-28-64-76 (1-400) x AB026980 (1-9264)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
Db 323 ATGGAAGTGAAGAACGCAGACCCCTACCGCTCTCTGACCTCCAGGGCGGACACGGAGCGC 382
QY 21 ArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 383 CGCTACACAGCTCATCCGCCGACAGCGAGGATGGCAAGATCAACCCCTAAATCTTACAGC 442
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 443 TCGAGTGAGACCTCAAAGCCTTCGACCAGGACTCCAGACTCGGCTTATGGCAGCCGCGTC 502
QY 61 LysAspIle**ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 503 AAAGACCTGGTGCACCATGAGGCGGACGAGTTCAGCAGACAAAGGCGCGGACTTTCTCTC 562
QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 563 AGAGACATGGCCTTCGGAGATCCCGTGC CGCCACATATGGGAGCGTACCGGACAGAAATG 622
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 623 GGCCTTCCCCACCGGACTACTCAGTGAGTGTGGCATCAGACGCTGACACAGAAACAGAC 682
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSer---ThrArgSerGly 139
Db 683 GGCATAATGTCCCCCGAGCACGCCGTCCGCTCTGCTATGCTCATCTGCCTCCTCTTCCCTGTCCGGC 742
QY 140 ArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHis 159
Db 743 CGCAGTTCTCCTGCTGTTTCAGCAGAGCCAACTCCAACCTTACGCTCACCGACACCGAGCAT 802
QY 160 GluAsnThrGluThrAspHisPro----- 167
Db 803 GAAACACTGAGAAATGTCCTCCTCTGCTATGCTCATCTGCCTCCTCTTCCCTGTCCGTCGAC 862
QY 167 ----- 167
Db 863 TCCCTTACCCTTCCCTTCCACGCGCAATCAGACCCAGGAAGTTGCTAGGTAAC 922
QY 167 ----- 167
Db 923 AGCGGGGCTCAGGCGCGGGACTCTGAGTCGGAAGATGAGTTTGGCCCCAATTCATTTC 982
QY 168 -----Gly 168
Db 983 TTAGTTAAACCGGCTCAGGGAACGTCTGCACTCCGGCGCGCGCTACTGCTAATGAGGGC 1042
QY 169 GlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHisAlaHisThr 188
Db 1043 TCGTTCAGAAATCACTACCGGTTGCGAACGCCCACTTACCCCTGTTTCACTCGCACTCC 1102
QY 189 ProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArg 208
Db 1103 CCCAGTCAGACCAACACACCGCTCCATCGGTTCCCTGAGCGCGAGCAACTACACCCAGCGC 1162
QY 209 SexAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGly 228
Db 1163 AGTAACCCGAGCGCGGACCCCGCATAGTTTCAGTCCCAATGAGGGTCCGACA---AGT 1219
QY 229 AlaGlnGluProAlaHisAlaGlnGluAsnTyrLeuLeuAsnSerAsnIleProLeuGlu 248
Db 1220 GCCCAGGACTCCAGCAGCGCGCCAGGACAACTGGCTTCTCAACAGCAACGTTCTCTTAGAG 1279
```

```
QY 249 ThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMet 268
Db 1280 ACCAGAAACATAGCAAGCAGACACTTCTAGAGACTTTGCAGGACAACTTTATAGAGATG 1339
QY 269 AspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys 288
Db 1340 GACATACTGGCGACAGCTCGCGCGGACGGCGCTTACACTGATGGACACTTCTTTTCAAG 1399
QY 289 ProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 1400 CCTGGCGGACCTCGCCTTTGTACTGTACCCACTTCACCTGGTTACCTTTGACGTCCAGC 1459
QY 309 ThrValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 1460 ACGGTGTACTCTCTCCGCCCGCCCTGTCCTGTCCTTCTACGGCCTGCCTTT 1519
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTyrLysCysAlaAlaLeuSerAlaIleVal 348
Db 1520 AGCCTAAAGAAACCCCTATAAGCACTGCAACTGGAAATGTGCTGCCCTCAGTGCCATCCTC 1579
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1580 ATCTCCGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1639
QY 369 AsnTyrHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSer 388
Db 1640 AACTGGCATTTGCAGCGGTGCAGAGGCGAGATATATCAGCTGACGGAGGACAAACACCACT 1699
QY 389 SerTrpProValProThrAspValSerLeuTyrPro 400
Db 1700 GGCCTACATCTGCCACAGATCTGGGGCTGCCACCT 1735

RESULT 12
AR270936
LOCUS AR270936 2387 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 11 from patent US 6500941.
ACCESSION AR270936
VERSION AR270936.1 GI:29702178
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2387)
AUTHORS Schaefer,G.M. and Sliwowski,M.
TITLE Gamma-herregulin
JOURNAL Patent: US 6500941-A 11 31-DEC-2002;
FEATURES Location/Qualifiers
source
1. 2387
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores: 4.21e-59 Length: 2387
Pred. No.: 1285.00 Matches: 235
Score: Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.27% Indels: 0
DB: 6 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AR270936 (1-2387)
QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCGCTGCAGAACCAACCGCGGCTCCGACGCCCGCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCCACACCCCAACCAACAGCACCGCGCTTCCATTAACTCCCTGAACCGGGGCAACTTC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCCGAGGAGCAACCCCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 180
```


QY	226	AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle	245
Db	181	GGCGGGCGGCCAGGAGCTGCCACGCCAGGAACTGGCTGCTCAACAGCAACATC	240
QY	246	ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu	265
Db	241	CCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTAGGGACATTGCAGGACAACCTC	300
QY	266	IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe	285
Db	301	ATTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGGCTTACAGTGACGGGCACTTC	360
QY	286	LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu	305
Db	361	CTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCACATCACAGGTACCCACTG	420
QY	306	ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg	325
Db	421	ACGTCCAGCACAGTGACTCTCTCCGCCCCGACCCCTGCCCGCAGCACCTTCGCCCGG	480
QY	326	ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer	345
Db	481	CCGGCCTTTAACTCAAGAACCCCTCCAAGTACTGTAACTGGAAGTGCGCAGCCCTGAGC	540
QY	346	AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu	365
Db	541	GCCATCGTCATCTCAGCCACTCTGGTCACTCGCTGGCATACTTTGTGGCCATGCACCTG	600
QY	366	PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp	385
Db	601	TTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTATGAGATCACGGAGGAC	660
QY	386	ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	661	ACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCCCTATACCCC	705
RESULT 13			
AB025412			
LOCUS	AB025412	8964 bp	mRNA linear ROD 08-MAY-1999
DEFINITION	Mus musculus mRNA for Ten-m3, complete cds.		
ACCESSION	AB025412		
VERSION	AB025412.1 GI:4760779		
KEYWORDS	Ten-m3.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (sites) Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T., Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.		
TITLE	Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues		
JOURNAL	J. Cell Biol. (1999) In press		
REFERENCE	2 (bases 1 to 8964)		
AUTHORS	Oohashi,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho, Okayama, Okayama 700-8252, Japan		
	(E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128, Fax: +81-86-222-7768)		
FEATURES	Location/Qualifiers		
source	1..8964		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="Balb/c"		
	/db_xref="taxon:10090"		
	/tissue_type="brain"		
	/dev_stage="adult"		
gene	1..8964		
	/gene="ten-m3"		
CDS	50..8197		

/gene="ten-m3"			
/codon_start=1			
/product="Ten-m3"			
/protein_id="BAA77398.1"			
/db_xref="GI:4760780"			
/translation="MDVKERRPYCSLTKSRRKERRYTNSSADNEECRVPTQKSYSSS			
ETLKAFDHDYSRLLYGNRVKDLVHREADEYTRQGNFTLRQLGVCESATRRGVAFCAE			
MGLPHRGYSISAGSDADTENEAVMSPEHAMRLWGRVKSGRSSCLSSRSNSALTLTDT			
EHENRSDSESEQPSNNPGQPTLOPLPPSHKHQHPAQHPSITSLNENSLTNRRNQSPAP			
PAALPAELQTTPTESVQLQDSWVLGSNVPLESRHFLFKTGTTPLTFSTATPGYTMASG			
SVSPPTPLPRNTLSRSFAFKFKSKYCSWRCTALCAVGSVLLAILLSYFIAMHLF			
GLNWHLQQTENDTFENGKVNSDTPTNTVSLPSGDNGLGFTHENNTIDSGLDIGH			
RAIQEVPPIFWRSOLFIDQPOFLKFENISLQDALIGVYGRKGLPPSHTYQDFVELLD			
GSRLIAREQRNLVESERAGQARSVSLHEAGFIQYLDSGIWHLAFYNDGKNPEQVSEFN			
TIVIESVVECPRNCHNGECVSGTCHCFPGFLGPDCSRACPVLCSGNGQYSKGRCLC			
FSGWKGTEDCVPTTQCIDPQCGRGICIMGSCACNSYKGENCEEADCLDPGCSNHGV			
CIHGECHCNPGWGSNCEILKTMADQCSGHGTYLQESGCTCDPNWTGPDCSNEICS			
VDCSHGVCMGGSCRCCEGTGPACNQRAChPRCAEHGTCKDGKCECSQGWNGEHCTI			
AHYLDKIVKEGCPGLCNSNGRCTLQDNGWHCVCPQGWGAGCDVAMETLCTDSKDN			
DGLIDCMDPDCCLOSCQNPYCRGLPDPODIISQSLQTPSQAAKSFYDRISFLIGS			
DSTHVLPGESPFNKSLASVIRGQVLTADGTPLIQNVNSFLHYSEYGYTITRODGMFDL			
VANGGASLTIVFERSPFLTQYHTVWIPWNVYVMDTLVMKKEENDIPSCDLSGFVRPS			
PIIVSSPLSTFFRSSPEDSPIPETQVLHEETTIPGTDLKLsyLSSRAAGYKSVLKIT			
MTQAVIPFNLMKVHLMAVVRGLFQKWPASPNLATYTFIMDKTDAYNQKVIYLGSEAVV			
SVGEYESCLDLTLWEKRTALQGYELDASNMGGWTLDKHHVLDVQNGILYKNGENQ			
FISQPPVSVSSIMGNRRRSISCPSCNGQADGNKLLAPVALACGIDGSLYVGDFNVVR			
RIPSGNVTSLVLELRNKFHSSNPAHRYYLATPVTGDLVYSDTNTRRIRYRPKSLTG			
AKDLTKNAEVVAGTGEQCLPFDEARCGDGKAVEATLSPKGMADKNGLIYFVDGTM			
IRKVDQNGIISTLLGSNDLTARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDM			
NVVLQITENRQVRIAAGRPMHCQVPGVEYYPVGKHAVQTLESATAIAVSYSLYTE			
TDEKKNIRIQVTTDGEISLVAGIPSECDCKNDANCDQCSGDGYAKDAKLNAPSSLA			
ASPDGTLYIADLGNIRIRAVSKNPLLNSMNFYEASPTDQELYIFDINGTHQYTVSL			
VTGDYLYNFYSYNDNDVATVDSNGNTLRIRRDPNMPVRVSPDNQVIWLTIGTNGC			
LKSMTAQGLELVLTFTYHGNSGLLATKSDETGWTTFDYDSEGRLTNTVFTPTGVVTNLH			
GDMDKAITVDIESSSREEDVSIITSNLSSIDSFTYMWQDLRNSYQIGYDGLRIFYAS			
GLDSHYQTEPHVLAGTANPTVAKRNMTLPGENGQNLVEWRFRKEAQQKVNVFGRKLR			
VNGRNLLSVDFTRTTEKIYDDHRKFLRLRIAYTDSGHTPLWLPSKLMVAVNTYSST			
GQIASIQRGTTSEKVDYDQGRIVSRVFADGKTYSYTYLEKSMVLLHLSQRQYIFEYD			
MMDRLSAITMPSVARHTMQTIRSIGYRNIYNPPESNASITDYNEEGLLLQTAFLGT			
SRRVLFKYRRQTRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIGPLI			
DRQIFRFSEDMVNAREFYSYDNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGK			
FGVIYDINQIIISTAVMTYTKHFDAGRIKEIQYEIYFRSLMWITIQYDNMGRVTKRE			
IKIGPFANTTKYAYEYVDVQLQTVYLNKIMRWRYNDLNGNLHLNPNSSSARLTPLR			
YDLDRITRLGDVQYRLDDEDFLRQRGTEIFEYSKGLLTRVYSKSGSWTVIYRYDGL			
GRRVSSKTSLGQHLQFPYADLTYPTRIITHVYNHSSSEITSLYDILQGLHFAMETSSGD			
EFYIASDNTGTPLAVSSNGLMLKQTYTAYGEIYFVDSNVDFQLVIGFHGGLYDELTK			
LIHFGERDYDILAGRWTTPDIEIWKRIQKDPAPFNLYMFRNNPNASKIHDVKDYITDV			
NSWLVTFGFHLHNAIPGFVPKFDLTPEPSYELVKSQWEDVPPIFGVQQQVARQAKAF			
LSLGKMAEVQVSRKAGAEQSWLWFATVKSLLIGKGMVAVSQGRVQTNVLNANEDCI			
KVAAVLNNAFYLENHLFTIEGKTDHYFIKTTTPESDLGTLRLTSGRKALENGINVTVS			
QSTTVNGNTRRFADVEMQFGALALHVRYGTMILDEEKARILEQARQARALARAWAREQQ			
RVRDGEEGARLWTEGEKRQLLSAGKVQGYDGYVYVLSVEYQPELADSAANNIQFLRQSEI			
GKR"			
ORIGIN			
Alignment Scores:			
Pred. No.:	1.67e-42	Length:	8964
Score:	982.00	Matches:	210
Percent Similarity:	62.65%	Conservative:	50
Best Local Similarity:	50.60%	Mismatches:	91
Query Match:	46.06%	Indels:	64
DB:	10	Gaps:	12
SEQ14-X-AT-28-64-76 (1-400) x AB025412 (1-8964)			
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArg---	ArgArgAspAlaGlu 19
Db	50	ATGGATGTGAAGGAACGACGGCCTTACTGCTCCTTGACCAAGAGCAGACGGGAAAGGAA	109
QY	20	ArgArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaPro---	GlnLysSer 38
Db	110	AGGCGCTATACAAATTCGTCGCGGACCAATGAGGAGTGTAGGGTCCCCCAGCAGAGTCC	169

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
|||
Db 170 TATAGTCCAGTGAAACCTTGAAGCTTTCGATCATGATTATTACGGCTGCTTTATGGA 229
|||
QY 58 SerArgVallysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
:::|||||
Db 230 AACAGAGTAAAGGATTGGTCCACAGAGAAGCCGACGAGTATACTAGACAAGACAGAAAT 289
|||
QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
|||
Db 290 TTTACCCTAAGGCAGTTAGGAGTGTGAATCCGCAACTCGAAGAGGAGTGGCATTTCTGT 349
|||
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
:::|||||
Db 350 GCGGAATGGGGTCCCTCACAGAGGTACTCCATCAGTGCAGGGTCAGATCGGATACG 409
|||
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
|||
Db 410 GAAACGAAAGCAGTGATGTCCCTGAGCATGCCATGAGACTTTGGGCAGGGGTCAAA 469
|||
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
|||
Db 470 TCGGGCCGAGTTCCTGCCTGTCAAGCCGCTCAACTCCGCCCTCACCTGACAGACACG 529
|||
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
|||
Db 530 GAGCACGAGAAGCAGGTCGGACAGTGAGCGGAGCAACCTTCAAAACAACCCAGGGCAACCC 589
|||
QY 175 ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
|||
Db 590 ACCCTGCAGCCTTTGCCGCCA-----TCCCAAGCAGCACCCGGCGGAGCATCAC--- 640
|||
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
|||
Db 641 CCGTCCATCACTCCCTCAATAGAAACTCCCTGACCAATAGAGGAACACAGATCCGGCC 700
|||
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
|||
Db 701 CCG-----CCGGCTGCTTTGCCCGCGAGCTGCAAAACCACA 736
|||
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
|||
Db 737 CCCGAGTCCGTCAGCTGCAGGACAGCTGGTCTTGGCAGTAATGTACCACTGGAAGC 796
|||
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
|||
Db 797 AGG----- 799
|||
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
|||
Db 800 -----CATTTCTCTATTCAAAACA 817
|||
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
|||
Db 818 GGGACAGGGACGACGCCACTGTTTCAGTACGGCAACCCGGGATACACAATGGCATCTGGC 877
|||
QY 309 ThrValTyrSerProProProArgProLeuProLeuPheThrPheAlaArgProAlaPhe 328
:::|||||
Db 878 TCTGTTTATTCTCCGCTACCCGGCCACTTCCTAGAAACACCCCTATCAAGAGTCTTTT 937
|||
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
|||
Db 938 AAATTCAAGAAGTCTTCAAAGTACTGCAGCTGGAGGTGCACCGCACTGTGTCTGTAGGG 997
|||
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
:::|||||
Db 998 GTCTCAGTGTCTCTGGCCATTCTCCTCTCCTATTTTATAGCAATGCATCTATTTGGCCTC 1057
|||
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
|||
Db 1058 AACTGGCACTTACAGCAGACCGGAAATGACACATTTCGAGATGGAAGTGAATTCGTGAC 1117
|||
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
|||

Db 1118 ACC-----GTGCCAACAAACACTGTATCGTTACCT 1147
|||
RESULT 14
AF195420 1476 bp mRNA linear PRI 27-JAN-2000
LOCUS Homo sapiens ODZ3 (ODZ3) mRNA, partial cds.
DEFINITION AF195420
ACCESSION AF195420.1 GI:6760372
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1476)
AUTHORS Ben-Zur,T., Feige,E., Motro,B. and Wides,R.
TITLE The mammalian Odz gene family: Homologs of a Drosophila pair rule gene with expression implying distinct yet overlapping developmental roles
JOURNAL Dev. Biol. 117, 107-120 (2000)
REFERENCE 2 (bases 1 to 1476)
AUTHORS Ben-Zur,T., Motro,B. and Wides,R.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel
FEATURES
source
1. .1476
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .>1476
/gene="ODZ3"
/note="similar to the Drosophila melanogaster pair rule odd Oz/Ten-m gene sequence"
/codon_start=1
/product="ODZ3"
/protein_id="AAF28318.1"
/db_xref="GI:6760373"
/translation="MDVKERRPYCSLTKSRREKERRYTNSSADNEECRVPTQKSYSSS
ETLKAFDHSRLLYGNRVKDLVHREADEFTROQNFTLRQLGVCEPATRRGLAFCAE
MGLPHRGYSISAGSDADTENEAVMSPEHAMRLWGRGVKSGRSSCISSRSNSALTLTDT
EHENKSDSENEQPASNQGSTLOPLPPSHKQHSQAHPHSITSLNRNSLTNRNQSPAP
PAALPAELQTTPTESVOLQDSWVLGSNVPLSRHFLKGTGTPLPSTATPGYTMASG
SVYSPTRPLPRNTLSRSFAFKFKSSKYCSWKCTALCAVGSVLLAILLSYFIAMHLF
GLNWQLQQTENDTFENGKVSNDTMTPTNTVSLPSGDNGKLGGFTQENNTIDSGELDIGR
RAIQEIPIPPGIFWRSQFLIDQPQFLKFNISLQKDALIGVYGRKKLPPSHTQ"

ORIGIN

Alignment Scores: 5.5e-43 Length: 1476
Pred. No.: 975.00 Matches: 210
Score: 62.41% Conservative: 49
Percent Similarity: 50.60% Mismatches: 92
Best Local Similarity: 45.73% Indels: 64
Query Match: 9 Gaps: 12
DB:

SEQ14-X-AT-28-64-76 (1-400) x AF195420 (1-1476)

QY 1 MetAspVallysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
|||
Db 151 ATGGATGTGAAGAACGACGGCCTTACTGCTCCCTGACCAAGAGCAGACGAGAAGGAA 210
|||
QY 20 ArgArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaPro---GlnLysSer 38
|||
Db 211 CGGCGCTACACAAATTCCTCCGACAGACAATGAGGAGTGCGGGGTACCCACACAGAGTCC 270
|||
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
|||
Db 271 TACAGTTCACGCGAGACATTGAAAGCTTTTGATCATGATTCTCGCGGCTGCTTTACGGC 330
|||

```
QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db 331 AACAGAGTGAAGGATTGGTTTCACAGAGAAGCAGACGAGTTCTACTAGACAAGACAGAAAT 390

QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCCCTAAGGCAGTTAGGAGTTTGTGAACCCAGCAACTCGAAGAGGACTGGCATTTTGT 450

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTGTATGTCCTCCAGAGCATGCCATGAGACTTTGGGGCAGGGGTCAAA 570

QY 138 SerGlyArgSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCCGCGAGCTCGCTGTCAAGTCGGTCCAACTCAGCCCTCACCCCTGACAGATACG 630

QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAGTCCGACAGTGAAGATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690

QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGCCGCT-----TCCCATAGCAGCACTTCGCACAGCATCAT--- 741

QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCCATCACTTCTCTCAACAGAAACTCCCTGACCAATAGAAGAAACCAGAGTCCGGCC 801

QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGGCTGCTTTGCCCCGCGAGCTGCAAAACCACA 837

QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCGAGTCCGTCCAGCTGCAGGACAGCTGGGTCTCTGGCAGTAATGTACCACTGGAAAGC 897

QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900

QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTCCTATTCAAACA 918

QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAACGCCACTGTTCTAGTACTGCAACCCAGGATACACAATGGCATCTGGC 978

QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTACCAACCTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTT 1038

QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAGAAGTCTTCAAAGTACTGTAGCTGGAATGCACCTGCACCTGTGTGCCGTAGGG 1098

QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCCTGGCAATACCTCCTGTCTTATTTTATAGCAATGCATCTCTTTGGCCTC 1158

QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGAAATGGAAGTGAATTCTGTAT 1218

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAAACACTGTGTCAATTACCT 1248
```

```
AX662355
LOCUS AX662355 8645 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 37 from Patent WO02062999.
ACCESSION AX662355
VERSION AX662355.1 GI:29163217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millet,I. and Macdougall,J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 37 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1..8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
```

```
Alignment Scores:
Pred. No.: 3.76e-42 Length: 8645
Score: 975.00 Matches: 210
Percent Similarity: 62.41% Conservative: 49
Best Local Similarity: 50.60% Mismatches: 92
Query Match: 45.73% Indels: 64
DB: 6 Gaps: 12
```

SEQ14-X-AT-28-64-76 (1-400) x AX662355 (1-8645)

```
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAGAACGCGAGGCTTACTGCTCCCTGACCAAGAGCAGACGAGAGAGGAA 210

QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGGCGCTACACAAATTCCTCCGACAGACAATGAGGAGTGCGGGTACCCACACAGAAAGTCC 270

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCACGCGAGACATGAAAGCTTTTGATCATGATTCCTCGCGCTGCTTTACGGC 330

QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db 331 AACAGAGTGAAGGATTGGTTTCACAGAGAAGCAGACGAGTTCACTAGACAAGACAGAAAT 390

QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCCCTAAGGCAGTTAGGAGTTTGTGAACCCAGCAACTCGAAGAGGACTGGCATTTTGT 450

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTGTATGTCCTCCAGAGCATGCCATGAGACTTTGGGGCAGGGGGTCAAA 570

QY 138 SerGlyArgSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCCGCGAGCTCGCTGTCAAGTCGGTCCAACTCAGCCCTCACCCCTGACAGATACG 630

QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAGTCCGACAGTGAAGATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690
```


QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGCCTGCT-----TCCCATAGAGCAGCACTCTGCACAGCATCAT--- 741

QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCTCAACAGAACTCCCTGACCAATAGAGAACCCAGAGTCCGGCC 801

QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CGGCTGCTTTGCCCGCCGAGCTGCAAAACCACA 837

QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCGAGTCCGTCAGCTGCAGGACAGCTGGGTCTTGGCAGTAATGTACCCTGGAAAGC 897

QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900

QY 270 IleLeuGlyAlaSerArgHisaspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTATTCAAAACA 918

QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAACGCCACTGTTCACTACTGCAACCCAGGATACACAATGGCATCTGGC 978

QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTACACCACTACTCGGCCACTACCTAGAAACACCCCTATCAAGAAGTGCTTT 1038

QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAGAAAGTCTTCAAGTACTGTAGCTGGAAATGCACCTGCTGTGTCGGTAGG 1098

QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCCTGGCAATACTCCTGTCTTATTATTATAGCAATGCATCTTTGGCCTC 1158

QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGATGGAAAGTGAATTTCTGAT 1218

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAAACACTGTGTCTATTACCT 1248

Search completed: August 14, 2004, 23:48:40
Job time : 5209 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 1255.76 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520
Perfect score: 372
Sequence: 1 TFWRSQVFIDHPVHLKFNVS.....RLLTQEARSLGTPRQSRGT 71

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29:	em_vi:*
30:	em_htg_hum:*
31:	em_htg_inv:*
32:	em_htg_other:*
33:	em_htg_mus:*
34:	em_htg_pln:*
35:	em_htg_rod:*
36:	em_htg_mam:*
37:	em_htg_vrt:*
38:	em_sy:*
39:	em_htgo_hum:*
40:	em_htgo_mus:*
41:	em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	372	100.0	1680	6	AR270935	AR270935 Sequence
2	372	100.0	2387	6	AR270936	AR270936 Sequence
3	372	100.0	3111	6	BD270887	BD270887 Method of
4	372	100.0	3111	6	AR270934	AR270934 Sequence
5	372	100.0	3111	6	BD140153	BD140153 Utilizati
6	372	100.0	3111	9	AF009227	AF009227 Homo sapi
7	372	100.0	8354	6	AX556500	AX556500 Sequence
8	372	100.0	8438	6	AX675551	AX675551 Sequence
9	372	100.0	8645	6	AX600210	AX600210 Sequence
10	353	94.9	8585	10	AB025413	AB025413 Mus muscu
11	347	93.3	9722	10	AF059485	AF059485 Mus muscu
12	277	74.5	9264	5	AB026980	AB026980 Danio rer
13	254	68.3	8473	6	AX662357	AX662357 Sequence
14	254	68.3	8487	6	AX662359	AX662359 Sequence
15	254	68.3	8645	6	AX662355	AX662355 Sequence
16	248.5	66.8	7816	10	AF195418	AF195418 Mus muscu
17	248.5	66.8	8816	5	AB026979	AB026979 Danio rer
18	248.5	66.8	8964	10	AB025412	AB025412 Mus muscu
19	234	62.9	8675	6	AX662353	AX662353 Sequence
20	232.5	62.5	173756	2	AC134948	AC134948 Rattus no
21	231.5	62.2	93474	2	AC136801	AC136801 Rattus no
22	231.5	62.2	245722	2	AC120288	AC120288 Rattus no
C 23	228	61.3	226256	10	AC073599	AC073599 Mus muscu
24	222.5	59.8	8118	5	GGA238613	AJ238613 Gallus ga
25	222	59.7	165812	2	AP001141	AP001141 Homo sapi
C 26	222	59.7	166063	9	AP002515	AP002515 Homo sapi
C 27	222	59.7	168308	2	AC084775	AC084775 Homo sapi
28	222	59.7	175464	2	AP002887	AP002887 Homo sapi
29	220.5	59.3	8373	10	AB025410	AB025410 Mus muscu
30	220	59.1	8297	9	AF100772	AF100772 Homo sapi
31	194.5	52.3	1430	6	AX250006	AX250006 Sequence
32	194.5	52.3	1431	6	AX058614	AX058614 Sequence
33	194.5	52.3	1727	6	AX714026	AX714026 Sequence
34	194.5	52.3	1727	9	AK056053	AK056053 Homo sapi
35	194.5	52.3	1743	6	AX133821	AX133821 Sequence
36	194.5	52.3	2496	5	GGA245711	AJ245711 Gallus ga
37	194.5	52.3	2496	6	AX250064	AX250064 Sequence
38	194.5	52.3	7713	9	HSM808325	BX648178 Homo sapi
39	194.5	52.3	8409	5	GGA279031	AJ279031 Gallus ga
40	194.5	52.3	8409	6	AX250068	AX250068 Sequence
41	194.5	52.3	8575	6	AX921803	AX921803 Sequence
42	194.5	52.3	8689	6	AX250067	AX250067 Sequence
43	194.5	52.3	8689	10	AF086607	AF086607 Rattus no
44	194.5	52.3	8797	6	AX250063	AX250063 Sequence
45	194.5	52.3	8797	6	AX250066	AX250066 Sequence

ALIGNMENTS

```
AR270935
LOCUS AR270935 1680 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 3 from patent US 6500941.
ACCESSION AR270935
VERSION AR270935.1 GI:29702177
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1680)
AUTHORS Schaefer,G.M. and Sliwkowski,M.
TITLE Gamma-herregulin
JOURNAL Patent: US 6500941-A 3 31-DEC-2002;
FEATURES
    source
        1..1680
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 6.57e-35 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x AR270935 (1-1680)
Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1407
Qy 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuThrGlnGluAlaArgSer 40
Db 1408 CTGGGAAGGCAGCCCTGGTGGCATTATGGCAGAAAGGCCCTCCTTCACATACA 1467
Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCTAACCCAGGAGCGCGGAGC 1527
Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1528 CTAGAGGGGACCCCGGCCAGTCTCGGGGAAC 1560
RESULT 2
AR270936
LOCUS AR270936 2387 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 11 from patent US 6500941.
ACCESSION AR270936
VERSION AR270936.1 GI:29702178
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2387)
AUTHORS Schaefer,G.M. and Sliwkowski,M.
TITLE Gamma-herregulin
JOURNAL Patent: US 6500941-A 11 31-DEC-2002;
FEATURES
    source
        1..2387
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 9.32e-35 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
```

```
US-10-029-020-14_COPY_450_520 (1-71) x AR270936 (1-2387)
Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 912
Qy 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 913 CTGGGAAGGCAGCCCTGGTGGCATTATGGCAGAAAGGCCCTCCTTCACATACA 972
Qy 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCTAACCCAGGAGCGCGGAGC 1032
Qy 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGGACCCCGGCCAGTCTCGGGGAAC 1065
RESULT 3
BD270887
LOCUS BD270887 3111 bp DNA linear PAT 17-JUL-2003
DEFINITION Method of promoting the proliferation of inner ear hairy cells
using ligand of HER2 receptor and/or HER3 receptor.
ACCESSION BD270887
VERSION BD270887.1 GI:33080655
KEYWORDS JP 2002529425-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3111)
AUTHORS Gao,W.Q.
TITLE Method of promoting the proliferation of inner ear hairy cells
using ligand of HER2 receptor and/or HER3 receptor
JOURNAL Patent: JP 2002529425-A 6 10-SEP-2002;
GENENTECH INC
COMMENT OS Homo sapiens (human)
PN JP 2002529425-A/6
PD 10-SEP-2002
PF 28-OCT-1999 JP 2000580655
PR 07-NOV-1998 US 60/107522
PI WEI QIANG GAO
PC A61K45/00,A61K38/00,A61K39/395,A61P27/16,C12N5/06//(C12N5/06,
C12R1:91),
PC A61K37/02,C12N5/00,(C12N5/00,C12R1:91)
CC Method of promoting the proliferation of
inner ear hairy cells
using
CC ligand of HER2 receptor and/or HER3 receptor
FH Key Location/Qualifiers
FT source 1..3111
FT /organism='Homo sapiens (human)'
FEATURES
    source
        Location/Qualifiers
            1..3111
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.21e-34 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x BD270887 (1-3111)
Qy 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1740
```


QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1741 CTGGGAAAGGCAGCCCTGGTTGGCATTATATGGCAGAAAGGCCTCCTCTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCGCCAGTCTCGGGAACT 1893

RESULT 4
AR270934
LOCUS AR270934 3111 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6500941.
ACCESSION AR270934
VERSION AR270934.1 GI:29702176
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3111)
AUTHORS Schaefer,G.M. and Sliwkowski,M.
TITLE Gamma-heregulin
JOURNAL Patent: US 6500941-A 1 31-DEC-2002;
FEATURES Location/Qualifiers
source 1..3111
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1.21e-34 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AR270934 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1741 CTGGGAAAGGCAGCCCTGGTTGGCATTATATGGCAGAAAGGCCTCCTCTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCGCCAGTCTCGGGAACT 1893

RESULT 5
BD140153
LOCUS BD140153 3111 bp DNA linear PAT 18-SEP-2002
DEFINITION Utilization of heregulin as epithelial cell growth factor.
ACCESSION BD140153
VERSION BD140153.1 GI:23235098
KEYWORDS JP 2002509076-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3111)
AUTHORS Sliwkowski,M. and Kern,J.A.
TITLE Utilization of heregulin as epithelial cell growth factor
JOURNAL Patent: JP 2002509076-A 6 26-MAR-2002;

GENENTECH INC, UNIVERSITY OF IOWA RESEARCH FOUNDATION
OS Homo sapiens (human)
PN JP 2002509076-A/6
PD 26-MAR-2002
PF 03-FEB-1999 JP 2000530226
PR 04-FEB-1998 US 09/020598
PI MARK SLIWKOWSKI,JEFFREY A KERN
PC A61K38/00,A61K35/12,A61K39/395,A61K39/395,A61P11/00,A61P43/00,
PC C07K14/485,
PC C07K16/22,C12N15/09,C12P21/08,A61K37/02,C12N15/00 CC
Utilization of heregulin as epithelial cell growth factor FH Key
Location/Qualifiers
FT source 1..3111
FT /organism='Homo sapiens (human)'.
Location/Qualifiers
1..3111
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.21e-34 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x BD140153 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1741 CTGGGAAAGGCAGCCCTGGTTGGCATTATATGGCAGAAAGGCCTCCTCTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCGCCAGTCTCGGGAACT 1893

RESULT 6
AF009227
LOCUS AF009227 3111 bp mRNA linear PRI 17-SEP-1997
DEFINITION Homo sapiens gamma-hereregulin mRNA, complete cds.
ACCESSION AF009227
VERSION AF009227.1 GI:2406643
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3111)
AUTHORS Schaefer,G., Fitzpatrick,V.D. and Sliwkowski,M.X.
TITLE Gamma-hereregulin: a novel heregulin isoform that is an autocrine growth factor for the human breast cancer cell line, MDA-MB-175
JOURNAL Oncogene 15 (12), 1385-1394 (1997)
MEDLINE 97472144
PUBMED 9333014
REFERENCE 2 (bases 1 to 3111)
AUTHORS Schaefer,G., Fitzpatrick,V.D. and Sliwkowski,M.X.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1997) Protein Chemistry, Genentech, Inc., 460 Point San Bruno Blvd, South San Francisco, CA 94080, USA
FEATURES Location/Qualifiers
source 1..3111
/organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22-p11"
/cell_line="MDA-MB-175"
/cell_type="breast cancer cell"
334..2640
/codon_start=1
/product="gamma-herregulin"
/protein_id="AAC51756.1"
/db_xref="GI:2406644"
/translation="MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSET
LKAYDQDARLAYGSRVKDIPVQEAEEFCRLTGANTFRLRELGLEEVTPPHGTLVRTDIGL
PHCGYSMGAGSDADMEADTVLSPEHPVRLWGRSTFSGRSSCLSSRANSNLTLTDTHEH
NTETDHPGGLQNHARLFTPPPPLSHAHTPNQHHAASINSLNRGNFTPRSNPSAPTDH
SLSGEPPAGGAQEPAAHQENWLLNSNIPLETRNLGKQPFGLTLDQNLIEMDILGASRH
DGAYSDGHFLFKPGGTSPLFCTTSPGYPLTSVTYSPPPRPLPRSTFARPAFNLKPPS
KYCNWKCAALSIAVISATLIVLLAYFVAMHLFGLNWLQPMEGOMYEITEDASSWPV
PTDVSLYPSGGTGLETPDRKGKGTTEGKPSFFPEDSFIDSGEIDVGRRASQKIPPGT
FWSQVFIIDHPVHLKENVSLGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQEAR
SLEGTQRQSGRTVPSSHETGFIQYLDSGIWHLAFYNDGKESEVVSFLTITATLPPRL
KEMKQSAAGSKLVLRCTESSEYSSLRFKWFKNGNELNRKNKPNQIKIQKPKGKSEL
RINKASLADSGEYMKVISKLGNDSSASANITIVESNELITGMPASTEGAYVSSSEPIR
ISVSTEGANTSSSTSTGTGTSHLVKCAEKEKTFVCVNGGECFMVKDLSNPSRYLCKCP
NEFTGDRCONYVMASFYSTSTPFLSLPE"
ORIGIN

Alignment Scores:
Pred. No.:      1.21e-34      Length:      3111
Score:          372.00        Matches:      71
Percent Similarity: 100.00%    Conservative:  0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    100.00%      Indels:       0
DB:              9           Gaps:        0

US-10-029-020-14_COPY_450_520 (1-71) x AF009227 (1-3111)

Qy      1  ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
      |||||||
Db      1681 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1740

Qy      21  LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
      |||||||
Db      1741 CTGGGAAAGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCCCTCCCTCCTTCACATACA 1800

Qy      41  GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
      |||||||
Db      1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1860

Qy      61  LeuGluGlyThrProArgGlnSerArgGlyThr 71
      |||||||
Db      1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 7
AX556500
LOCUS      AX556500      8354 bp      DNA      linear      PAT 27-NOV-2002
DEFINITION      Sequence 13 from Patent WO02057453.
ACCESSION      AX556500
VERSION      AX556500.1  GI:25899736
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,
Smithson,G., Li,L. and Ji,W.
Polypetides and nucleic acids encoding same
Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
LOCATION/Qualifiers

ORIGIN

Alignment Scores:
Pred. No.:      3.27e-34      Length:      8438
Score:          372.00        Matches:      71
Percent Similarity: 100.00%    Conservative:  0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    100.00%      Indels:       0
DB:              6           Gaps:        0

US-10-029-020-14_COPY_450_520 (1-71) x AX675551 (1-8438)

Qy      1  ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
      |||||||
Db      1382 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1441

Qy      21  LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
      |||||||
Db      1442 CTGGGAAAGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCCCTCCCTCCTTCACATACA 1501

Qy      41  GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
      |||||||
Db      1502 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1561

RESULT 8
AX675551
LOCUS      AX675551      8438 bp      DNA      linear      PAT 27-MAR-2003
DEFINITION      Sequence 1 from Patent WO02055704.
ACCESSION      AX675551
VERSION      AX675551.1  GI:29333552
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
Stone,D.J.
Proteins, polynucleotides encoding them and methods of using the
same
Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
LOCATION/Qualifiers

ORIGIN

Alignment Scores:
Pred. No.:      3.27e-34      Length:      8438
Score:          372.00        Matches:      71
Percent Similarity: 100.00%    Conservative:  0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    100.00%      Indels:       0
DB:              6           Gaps:        0

US-10-029-020-14_COPY_450_520 (1-71) x AX675551 (1-8438)

Qy      1  ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
      |||||||
Db      1562 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1594
```

```

source
1..8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.:      3.24e-34      Length:      8354
Score:          372.00        Matches:      71
Percent Similarity: 100.00%    Conservative:  0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    100.00%      Indels:       0
DB:              6           Gaps:        0

US-10-029-020-14_COPY_450_520 (1-71) x AX556500 (1-8354)

Qy      1  ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
      |||||||
Db      1382 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1441

Qy      21  LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
      |||||||
Db      1442 CTGGGAAAGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCCCTCCCTCCTTCACATACA 1501

Qy      41  GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
      |||||||
Db      1502 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1561

RESULT 8
AX675551
LOCUS      AX675551      8438 bp      DNA      linear      PAT 27-MAR-2003
DEFINITION      Sequence 1 from Patent WO02055704.
ACCESSION      AX675551
VERSION      AX675551.1  GI:29333552
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1
AUTHORS      Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
Stone,D.J.
Proteins, polynucleotides encoding them and methods of using the
same
Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
LOCATION/Qualifiers

ORIGIN

Alignment Scores:
Pred. No.:      3.27e-34      Length:      8438
Score:          372.00        Matches:      71
Percent Similarity: 100.00%    Conservative:  0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    100.00%      Indels:       0
DB:              6           Gaps:        0

US-10-029-020-14_COPY_450_520 (1-71) x AX675551 (1-8438)

Qy      1  ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
      |||||||
Db      1562 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1594
```

Db 1358 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCCTGTGCATCTGAAATTCAATGTGTCT 1417
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1418 CTGGAAAGGCAGCCCTGGTGGCATTTATGGCAGAAAAGGCCTCCCTCCTTCACATACA 1477
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1478 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCTTAACCCAGGAGGCGCGGAGC 1537
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1538 CTAGAGGGGACCCCGCCAGTCTCGGGGAAC 1570
RESULT 9
AX600210
LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Walia,N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="caxon:9606"
/note="Incyte ID No: 7488573CH1"
ORIGIN
Alignment Scores:
Pred. No.: 3.35e-34 Length: 8645
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6
US-10-029-020-14_COPY_450_520 (1-71) x AX600210 (1-8645)
QY 1 ThrPheTyrArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1452 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCCTGTGCATCTGAAATTCAATGTGTCT 1511
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1512 CTGGAAAGGCAGCCCTGGTGGCATTTATGGCAGAAAAGGCCTCCCTCCTTCACATACA 1571
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1572 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCTTAACCCAGGAGGCGCGGAGC 1631
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1632 CTAGAGGGGACCCCGCCAGTCTCGGGGAAC 1664
RESULT 10
AB025413
LOCUS AB025413 8585 bp mRNA linear ROD 08-MAY-1999
DEFINITION Mus musculus mRNA for Ten-m4, complete cds.
ACCESSION AB025413

VERSION AB025413.1 GI:4760781
KEYWORDS Ten-m4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Ohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
JOURNAL J. Cell Biol. (1999) In press
REFERENCE 2 (bases 1 to 8585)
AUTHORS Ohashi,T.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128,
Fax:+81-86-222-7768)
FEATURES
source Location/Qualifiers
1. .8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1. .8585
/gene="ten-m4"
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translation="MDVKERKPYRSLTRRDAERRYTSSSADSEEGKGPKQKYSSET
LKAYDQARLAYGSRVKOMVPQEAEEFCRTGTNTLRELGLGEMTPPHGTLYRTDIGL
PHCGYSMGASSDADLEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTEHE
NTETDHPSSLQNHPRLRTPPPLPHAHTPNQHHAAASINSLNRGNFTPRSNPSPAPT
SLSGEPPAGSAQEPHTAQNWLNSNIPLERNLGKQPFGLTLQDNLNLIEMDILSASRH
DGAYSDGHFLFKPGGTSPLFCTTSPGYPLTSTVSPRPRLPRSRPAFNLKKPS
KYCNWKCAALSAILISATLIVLLAYFVAMHFLGLNWHLQPMEGQMVEITEDTASSW
PVPTDVSLYPSGGTLETPDRKGKGAEGKPSLFPEDSFIDSGEIDVGRRASQKIPP
GTFWRSQVFIIDHPVHLKFNVSGLKAAVLGYGRKGLPPSHTOQDFVELDGRRLTQE
ARSLGEPQROSRGPVPPSSHETGFIQYLDSGIWHLAFYNDKSESEVFLTTAIESVD
NCPSNCYNGDCISGTCHEFLGFLGPCGRASCPLVCSGNGQYMKGRCLCHSGWKGA
CDVPTNQCIDVACSSHGTCIMGTICNPGYKGSECEVEDCMDPTCSSRGVCVRGEC
SVWGGTNCETPRATCLDQC SGHGTFLDTGLCNCDPSWTGHDCSEI CAADCGGHV
CVGGTCRCEDGMWGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHCTIAHYLD
KEGCPGLCNGNRCITLDLNGWHCVCLGWRGTGCDTSMETCGDGDNDGDGLVDCMD
PDCLQLCHVNPLCLGSPDLDI IQETAPVSQQNLNPFYDRIKFLVGRDSTHSIPG
ENPFDGHCACVIRGQVMTSDGTPLVGNISFINNPLFGYTTISRQDGSFDLVTNGGSI
ILRFERAPFTQETHLWLPWDRFFVMEIVMRHENEIPSCDLSNEAPNPVSPSPL
TSFASSCAEKGPVPEIQALQEEIVIAGCKMRLSYLSSRTPGYKSVLRISLTHPTIPF
NLMKVHLMVAVEGRLFRKWFAAAPDLSYFIWDKTDVYNQKVFGESEAFVSVGYEYES
CPDLILWEKRTAVLQGYEIDASKLGGWSLDKHHALNIQSGILHKNGENQFVSQQPPV
IGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGLYVGDEFYIRRIFFSGNV
TNILEMRNKDPRHSHSPAHHYYLATDPMGSAVFLSDTNSRRRVFKVKSTTVVKDLVKN
EVAGTGDCQLPFDDTRCGDGGKATEATLTPRGITVDKFGLIYFVDGTMIRRVQNG
IISTLLGSNDLTSARPLSCDSVMEISQVLEWPTDLAINPMDNSLYLDNNVVLQISE
NHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKIN
RIRQVTTSGEISLVAGAPSGCDCKNDANCDFSGDYAKADAKLNPSSLAVCADGEL
YVADLGNIRIRFIRKNKPFNLNTQNMVEISPIDQELYFDTSGHLYTQSLVCTDGLY
NFTYTGDDITHITDNNGMNVNRDSTGMPLWLVPDQGVYVWVTMGTNSALRSVTTQ
GHELAMTYHNSGLLATKSNENGWTFYEYDSFGRLTNVTFTPGQVSSFRSDTDSV
HVQVETSSKDDVTITTNLSAGFAYTLIQDQVRNSYIGADGSLRLLLANGMEVALQT
EPHLLAGTVNPTVGKRNVTLPIDNGLNLVEWRQKEAQGVTFGRRLRVHNRNLLS
LDFDRVTRTEKIYDDHRKFTLRILYDQGRPSLWSPSSRLNGVNVYSPGGHIAGIOR
GIMSERMEYDQAGRITSRIFADGKMWSYTYLEKSMVLHLHSQRQYIPEFDKNDRLSSV
TMPNVARQTLLETIRSVGYRNIYQPPEGNASVIQDFTEDGHLHLHTFYLGTRRRVIKY
GKLSKLAETLYDTTKVSTFYDETAGMLKTVNLQNEGFTCTIRYRQIGPLIDRQIFRFT
EEGMVNARFDYDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEFGKFGVIYYDI

NQIIITAVMTHTKHFDAYGRMKEVQYEIFRSLMYWMTVQYDNNMGRVVKKELKVGYPYAN
TTRYSEYDADGQLQTVSINDKPLWRSYDLNGNLHLLSPGNSARLTPLRYDLDRIT
RLGDVQYKMDDEGLRQGGDVFEYNSAGLLIKAYNRASGWSVRYRGLGRRVSSKS
SHSHLQFFYADLTNPTKVTHLYNHSSEITSLYYDLQHLFAMELSSGDEFYIACDN
IGTPLAVFSGTGLMIKQILYIYAYGEIYMDTNPNFQIIGYHGLYDPLTKLVHMGRRD
YDVLAGRWTSPDHELWKRLSSNSIVPFHLYMFKNNPISNSQDIKCFMTDVNSWLLTF
GFQLHNVIPGYPKPDTDAMEPSYELVHTQMTQEWDNKSILGVQCEVQQLKAFVTL
ERFDQLYGSTITSCQAPETKKFASSGSIFGKGVKFAKDGKRVTTDIIISVANEDGRRRI
AAILNNAHYLENLHFTIDGVDTHYFVKPGPSEGDLAAILGLSGRRRTLENGVNTVSQI
NTMLSGRRTRYTDIQLQYRALCLNTRYGTTVDEKVRVLELARQAVRQAWAREQORL
REGEGLRAWTDGEKQQVLTGRVQGYDGFVTISVEQYPELSDSANNIHFMRQSEMGR
R"

ORIGIN

Alignment Scores:				
Pred. No.:	6.28e-32	Length:	8585	
Score:	353.00	Matches:	68	
Percent Similarity:	97.14%	Conservative:	0	
Best Local Similarity:	97.14%	Mismatches:	2	
Query Match:	94.89%	Indels:	0	
DB:	10	Gaps:	0	

US-10-029-020-14_COPY_450_520 (1-71) x AB025413 (1-8585)

QY	1	ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer	20
Db	1541	ACGTTCTGGAGATCTCAGGTGTTTCATAGACCACCTGTACACCTGAAGTTCAATGTGCT	1600
QY	21	LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr	40
Db	1601	CTGGGAAAGCAGCTCTGTTGGCATTATGGCAGAAAAGGCCTTCCTCCTCCCATCT	1660
QY	41	GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer	60
Db	1661	CAGTTTGACTTTGTGGAGCTCTGGATGGAAGAGGCTCTAACCCAGAGGCAAGGAGC	1720
QY	61	LeuGluGlyThrProArgGlnSerArgGly	70
Db	1721	CTGGAGGGTCCTCAGCGCCAATCACGGGGC	1750

RESULT 11
AF059485
LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998
DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 9722)
AUTHORS Wang,X.-Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P., Zinszner,H. and Ron,D.
TITLE Identification of novel stress-induced genes downstream of chop
JOURNAL EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE 98315054
PUBMED 9649432
REFERENCE 2 (bases 1 to 9722)
AUTHORS Wang,X.-Z. and Ron,D.
TITLE Direct Submission
JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University Medical Center, 550 First Ave., New York, NY 10016, USA

FEATURES

source
1. .9722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1. .9722
/gene="Doc4"

CDS

/note="downstream of CHOP 4"
583. .9060
/gene="Doc4"
/note="similar to Drosophila melanogaster tenm/odz and human gamma-hergulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translation="MDVKERKPYRSLRRRDAERRYTSSSADSEEGKGPQKSYSSSET LKAYDQARLAYGSRVKDMVPOEAEEFCRTGNTFTLRELGLGEMTPPHGTLVRTDIGL PHCGYSMGASSDADLEADTVLSPHPVRLWGRSTRSGRSSCLSSRANSNLITLDTHEH NTETGAPLHCSSASSTPIEQSPSPPPSPPANESQRRLLGNGVAQPTPDSSEEEFVPN SFLVKGSASLGAANDHPSSLQNHPRLRTPPPLPHAHTPNQHHAAASINSLNRGNFT PRSNPSPAPTDHLSLGEPPAGSAQEPTHAQDNWVLNSKIPVETRNLGKQFLGTWQDN LIEMDIFASRRDGAYSDBGHFFKPGGTSPLFCTTSPGYPLTSTSVSPRPRLPRST FSRPAFNLKPKSKYCNWKCAALSAILISATLVILLAYFVAMHLFGLNHLQPMEGQMQ MYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGAAGKPSLFPEDSFIDSGEI DVGRRASQKIPPGTFWRSQVFIHPVHLKFNVSGLKAALVGIYGRKGLPSPSTQLDFV ELLDGRRLLTQEARSLGEPQORSGRPVPPSSHETGFIQYLDSGIWHLAFYNDGKESEV VSFLTTAIESVDNCPSCYNGDCISGTCFCFLGFLGPDGCRASCPVLCSGNGQYMKG RCLCHSGWKGAECDDVTNQCIDVACSSHGTCIMGTICINPGYKGESCEEVDCMDPTCS SRGVCVRGECHSVGWGNTCETPRATCLDQCSHGTFLPDTGLCNCDPSTWGHDCS1 EICAADCGGHGVCVGTCCRCEDGWMGAACDQACRACHPRCAEHGTCRDCKECTPGWNGE HCTIEGCPGLCNGNRCCTDLNGWHCVCOLGWRGTCDTSMETCGCGDKNDGDGLVD CMDPDCCLQPLCHVNPCLCLGSPDPLDIIQETQAPVSOQNLNSFYDRKFLVGRDSTHS IPGENPEFDGGHACVIRGQVMTSDGTPLVGNISFINNPLFGYISRDQGSFDFLVTNGG ISIIILRFERAPFITQEHTLWLPWDRFFVMTIIVMRHEENEIIPSRDLSNFAFPNPVVSF SPLTSFASSCAEKGPVPEIQALQEEIIVAGCKMRLSYLSRTPGYSVVRISLTHPT IPFNLKVMHLMVAVEGRFLRKWFAPADLSYFIDWKTDVYNQKVFGLSEAFVSGYE YESCPDLILWEKRTAVLQGYEIDASKLGWSLDKHALNIQSGILHKNGENQFVSQQ PPVIGSIMNGRRRSISCPSCNGLADGNKLAPVALTCGSDGLYVGFNIRIRFIPS GNVNIIEMSHSPAHHYYLATDPMGAVFLSDTNSRRRVFKVKSTTVVXDLVKNSEVVA GTDQCLPFDDTRCGDGGKATEATLTNPRGITYDKFGLIYFVDGTMIRRVQDNGIIST LLGSDNLTARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYVLDDNNVVLQISENHQV RIVAGRPMHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLYIAETDEKINRIQ VTTSGEISLVAGAPSGCDCKNDANCDGDDGYAKDAKLNTPSSLAVCADGELYVAD LGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFDTSKGHLYTQSLPTGYLYNFY TGDDGDIITHITDNNGNMNVNRDSTGMPLWLVPDGGVYVMTGMTNSALRSVTTQGHF AMTYHGNSGLLATKSNENGWTTTFEYDSFGLTNVTFTPTQGVSSFRSDTSSVHVQV ETSSKDDVTITHLSSGSFATLLQDQVRNSYIIGADSLRLLLANGMEVALQTEPHL LAGTVNPTVGKRVNTPIDNGLNLVEWRQKEQARQVTVFGPRLRVHNRNLLSLDFD RVTRTEKIYDDHRKFTLRILYDQAGRPSFWSPPSLRNGVNVTVSPGGHAGTQRGIMS ERMEYDQAGRITSRIFADGKMWSYTYLEKSMVLHLHSRQYIIFEEDKNDRLSSVTMPN VARQTLTIRSVGYRNIYQPPEGNASVIQDFTEDGHLHTFYLGTGRRVYKYKGLS KLAETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCIYRQIGPLIDRQIFRFTEEGM VNAREFDYVNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVYDINQII TTAVMTHSKHFDAYGRMKEVQYETFRSLMYWMTVQYDNNMGRVVKKELKVGYPYANTRY SYEYDADGQLQTVSINDKPLWRSYDLNGNLHLLSPGNSARLTPLRYDLDRITRLGD VQYKXMDGSLRQGGDVFEYNSAGLLIKAYNRASGWSVRYRYDGLGRRVSSKSSHSH HLQFFYADLTNPTKVTHLYSHSSSEITSLYYDLQHLFAMELSSGDEFYIACDNIGTP LAVFSGTGLMIKQILYIYAYGEIYMDTNPNFQIIGYHGLYDPLTKLVHMGRRDYDVL AGRWTSPPDHELWKRLSSNSIVPFHLYMFKNNPISNSQDIKCFMTDVNSWLLTFGFQL HNVI PGYPKPDTDAMEPSYELVHTQMTQEWDNKSILGVQCEVQQLKAFVTLERFD QLYGSTITSCQAPETKKFASSGSIFGKGVKFAKDGRTVTDIIISVANEDGRRRIAIL NNAHYLENLHFTIDGVDTHYFVKPGPSEGDLAAILGLSGRRRTLENGVNVTVSQINTML SGRTRYTDIQLQYRALCLNTRYGTTVDEKVRVLELARQAVRQAWAREQORLREGE EGLRAWTDGEKQQVLTGRVQGYDGFVTISVEQYPELSDSANNIHFMRQSEMGRR"

ORIGIN

Alignment Scores:

Pred. No.:	3.72e-31	Length:	9722
Score:	347.00	Matches:	67
Percent Similarity:	95.71%	Conservative:	0
Best Local Similarity:	95.71%	Mismatches:	3
Query Match:	93.28%	Indels:	0
DB:	10	Gaps:	0

US-10-029-020-14_COPY_450_520 (1-71) x AF059485 (1-9722)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS

Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X., Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K., Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A., Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F., Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G., Millet,I. and Macdougall,J.R.

TITLE

PROTEINS AND NUCLEIC ACIDS ENCODING SAME

JOURNAL

PATENT: WO 02062999-A 39 15-AUG-2002;

CURAGEN CORPORATION (US)

LOCATION/QUALIFIERS

1. .8473

/ORGANISM="Homo sapiens"

/MOL_TYPE="unassigned DNA"

/DB_XREF="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 4.47e-20 Length: 8473

Score: 254.00 Matches: 48

Percent Similarity: 80.88% Conservative: 7

Best Local Similarity: 70.59% Mismatches: 13

Query Match: 68.28% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX662357 (1-8473)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21

DB 1185 TTCTGGAGATCACAGCTCTTCATGTGATCAGCCACAGTTTCTTAAATTCAATATCTCTCTT 1244

QY 22 GlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuThrGlnGluAlaArgSerLeu 41

DB 1245 CAGAAGGATGCATTGATTGGAGTATATGGCCGGAAGGCTTACCGCTTCCCATACTCAG 1304

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61

DB 1305 TATGACTTCGTGGAGCTCTCTGGATGGCAGCAGGCTGATGCCAGAGCAGCGGAGCCTG 1364

QY 62 GluGlyThrProArgGlnSerArg 69

DB 1365 CTTGAGACGGAGAGCGCGGCGG 1388

RESULT 14

AX662359

LOCUS

AX662359 Sequence 41 from Patent WO02062999.

ACCESSION

AX662359.1 GI:29163219

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS

Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X., Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K., Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A., Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F., Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G., Millet,I. and Macdougall,J.R.

TITLE

PROTEINS AND NUCLEIC ACIDS ENCODING SAME

JOURNAL

PATENT: WO 02062999-A 41 15-AUG-2002;

CURAGEN CORPORATION (US)

LOCATION/QUALIFIERS

1. .8487

/ORGANISM="Homo sapiens"

/MOL_TYPE="unassigned DNA"

/DB_XREF="taxon:9606"

ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS

Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X., Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K., Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A., Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F., Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G., Millet,I. and Macdougall,J.R.

TITLE

PROTEINS AND NUCLEIC ACIDS ENCODING SAME

JOURNAL

PATENT: WO 02062999-A 37 15-AUG-2002;

CURAGEN CORPORATION (US)

LOCATION/QUALIFIERS

1. .8645

/ORGANISM="Homo sapiens"

/MOL_TYPE="unassigned DNA"

/DB_XREF="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 4.56e-20 Length: 8645

Score: 254.00 Matches: 48

Percent Similarity: 80.88% Conservative: 7

Best Local Similarity: 70.59% Mismatches: 13

Query Match: 68.28% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX662355 (1-8645)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21

DB 1357 TTCTGGAGATCACAGCTCTTCATGTGATCAGCCACAGTTTCTTAAATTCAATATCTCTCTT 1416

QY 22 GlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProProSerHisThrGln 41

DB 1417 CAGAAGGATGCATTGATTGGAGTATATGGCCGGAAGGCTTACCGCTTCCCATACTCAG 1476

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61

Alignment Scores:

Pred. No.: 4.48e-20 Length: 8487

Score: 254.00 Matches: 48

Percent Similarity: 80.88% Conservative: 7

Best Local Similarity: 70.59% Mismatches: 13

Query Match: 68.28% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX662359 (1-8487)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21

DB 1226 TTCTGGAGATCACAGCTCTTCATGTGATCAGCCACAGTTTCTTAAATTCAATATCTCTCTT 1285

QY 22 GlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProProSerHisThrGln 41

DB 1286 CAGAAGGATGCATTGATTGGAGTATATGGCCGGAAGGCTTACCGCTTCCCATACTCAG 1345

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61

DB 1346 TATGACTTCGTGGAGCTCTCTGGATGGCAGCAGGCTGATGCCAGAGCAGCGGAGCCTG 1405

QY 62 GluGlyThrProArgGlnSerArg 69

DB 1406 CTTGAGACGGAGAGCGCGGCGG 1429

RESULT 15

AX662355

LOCUS

AX662355 Sequence 37 from Patent WO02062999.

ACCESSION

AX662355.1 GI:29163217

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

REFERENCE

AUTHORS

Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X., Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K., Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A., Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F., Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G., Millet,I. and Macdougall,J.R.

TITLE

PROTEINS AND NUCLEIC ACIDS ENCODING SAME

JOURNAL

PATENT: WO 02062999-A 37 15-AUG-2002;

CURAGEN CORPORATION (US)

LOCATION/QUALIFIERS

1. .8645

/ORGANISM="Homo sapiens"

/MOL_TYPE="unassigned DNA"

/DB_XREF="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 4.56e-20 Length: 8645

Score: 254.00 Matches: 48

Percent Similarity: 80.88% Conservative: 7

Best Local Similarity: 70.59% Mismatches: 13

Query Match: 68.28% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AX662355 (1-8645)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21

DB 1357 TTCTGGAGATCACAGCTCTTCATGTGATCAGCCACAGTTTCTTAAATTCAATATCTCTCTT 1416

QY 22 GlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProProSerHisThrGln 41

DB 1417 CAGAAGGATGCATTGATTGGAGTATATGGCCGGAAGGCTTACCGCTTCCCATACTCAG 1476

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61

Db	1477	TATGACTTCGTGGAGCTCCTGGATGGCAGCAGGCTGATTGCCAGAGACGCGGAGCCTG	1536
QY	62	GluGlyThrProArgGlnSerArg	69
Db	1537	CTTGAGACGGAGAGAGCCCGGCGG	1560

Search completed: August 14, 2004, 11:51:25
Job time : 1270.76 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 21.0053 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520

Perfect score: 372
Sequence: 1 TFWRSQVFIDHPVHLKENVS.....RLLTQEARSLGTPRQSRGT 71

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10029020/runat 06082004 112217 29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN 1 1 258 @runat 06082004 112217 29301 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCFUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372	100.0	1680	3	US-08-891-845-3 Sequence 3, Appli
2	372	100.0	1680	4	US-09-514-573-3 Sequence 3, Appli
3	372	100.0	2387	3	US-08-891-845-11 Sequence 11, Appl
4	372	100.0	2387	4	US-09-514-573-11 Sequence 11, Appl
5	372	100.0	3111	3	US-08-891-845-1 Sequence 1, Appli
6	372	100.0	3111	4	US-09-514-573-1 Sequence 1, Appli
C 7	69	18.5	1083	4	US-09-252-991A-11841 Sequence 11841, A
8	69	18.5	1407	4	US-09-252-991A-11543 Sequence 11543, A
C 9	69	18.5	2304	4	US-09-252-991A-11691 Sequence 11691, A
10	67.5	18.1	523	4	US-09-669-751-81 Sequence 81, Appl
C 11	66.5	17.9	1272	4	US-09-252-991A-14082 Sequence 14082, A
C 12	66.5	17.9	1329	4	US-09-252-991A-14036 Sequence 14036, A

13	66.5	17.9	1620	4	US-09-252-991A-14186 Sequence 14186, A
C 14	66	17.7	3835	4	US-09-566-921-120 Sequence 120, App
15	65	17.5	945	4	US-09-584-568C-3 Sequence 3, Appli
16	65	17.5	1332	4	US-09-584-568C-1 Sequence 1, Appli
17	65	17.5	1426	4	US-09-620-312D-967 Sequence 967, App
18	63.5	17.1	513	4	US-09-252-991A-1256 Sequence 1256, Ap
19	62.5	16.8	1233	4	US-09-252-991A-3735 Sequence 3735, Ap
C 20	62.5	16.8	1491	4	US-09-252-991A-3636 Sequence 3636, Ap
C 21	62.5	16.8	1812	4	US-09-732-615-27 Sequence 27, Appl
C 22	62.5	16.8	2242	3	US-09-400-742-1 Sequence 1, Appli
C 23	62.5	16.8	2242	3	US-08-618-651A-1 Sequence 1, Appli
C 24	62.5	16.8	2242	4	US-09-215-252-1 Sequence 1, Appli
C 25	62.5	16.8	2242	4	US-09-970-989A-1 Sequence 1, Appli
26	62.5	16.8	3301	4	US-09-148-545-66 Sequence 66, Appl
27	62.5	16.8	4884	4	US-09-252-991A-3824 Sequence 3824, Ap
28	62	16.7	432	4	US-08-956-171E-4061 Sequence 4061, Ap
C 29	62	16.7	480	4	US-09-252-991A-5103 Sequence 5103, Ap
C 30	62	16.7	801	4	US-09-252-991A-5129 Sequence 5129, Ap
31	62	16.7	1557	4	US-09-252-991A-5090 Sequence 5090, Ap
32	62	16.7	2855	2	US-08-852-153-1 Sequence 1, Appli
33	62	16.7	3255	2	US-08-852-153-5 Sequence 5, Appli
34	62	16.7	3327	2	US-08-852-153-3 Sequence 3, Appli
C 35	62	16.7	22243	4	US-08-956-171E-164 Sequence 164, App
C 36	61.5	16.5	1080	4	US-09-252-991A-15937 Sequence 15937, A
C 37	61.5	16.5	2721	4	US-09-252-991A-16144 Sequence 16144, A
38	61.5	16.5	11601	2	US-08-222-617A-3 Sequence 3, Appli
39	61.5	16.5	11601	2	US-08-222-617A-24 Sequence 24, Appl
C 40	61	16.4	4403765	3	US-09-103-840A-2 Sequence 2, Appli
C 41	61	16.4	4411529	3	US-09-103-840A-1 Sequence 1, Appli
42	60.5	16.3	309	4	US-09-252-991A-1165 Sequence 1165, Ap
C 43	60.5	16.3	1284	4	US-09-252-991A-13863 Sequence 13863, A
C 44	60.5	16.3	1293	4	US-09-252-991A-13668 Sequence 13668, A
45	60.5	16.3	1680	4	US-09-252-991A-7909 Sequence 7909, Ap

ALIGNMENTS

RESULT 1

US-08-891-845-3
; Sequence 3, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-891-845-3

Alignment Scores:
Pred. No.: 1.3e-43 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-08-891-845-3 (1-1680)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGTCATCTGAAATTCAATGTGCT 1407
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1408 CTGGAAAGGCAGCCCTGGTGGCATTTATGGCAGAAAGGCCTCCCTTCACATACA 1467
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1527
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1528 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1560

RESULT 2

US-09-514-573-3
; Sequence 3, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-3

Alignment Scores:
Pred. No.: 1.3e-43 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-514-573-3 (1-1680)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGTCATCTGAAATTCAATGTGCT 1407
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1408 CTGGAAAGGCAGCCCTGGTGGCATTTATGGCAGAAAGGCCTCCCTTCACATACA 1467
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1527
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1528 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1560

RESULT 3

US-08-891-845-11
; Sequence 11, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid

```

; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-891-845-11

Alignment Scores:
Pred. No.: 2.le-43 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-08-891-845-11 (1-2387)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGCATCTGAAATTCATGTGTCT 912
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 913 CTGGGAAAGGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCCCTCCCTTCACATACA 972
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAGC 1032
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1065

RESULT 4
US-09-514-573-11
; Sequence 11, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear

; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-891-845-11

Alignment Scores:
Pred. No.: 2.le-43 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-08-891-845-11 (1-2387)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGCATCTGAAATTCATGTGTCT 912
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 913 CTGGGAAAGGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCCCTCCCTTCACATACA 972
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAGC 1032
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1065

RESULT 5
US-08-891-845-1
; Sequence 1, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-891-845-1

```

Alignment Scores:
Pred. No.: 3.01e-43 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-08-891-845-1 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGAAAGGCAGCCCTGGTTGGCATTTATGGCAGAAAAGGCTCCCTTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGAGGCTCCTAAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1893

RESULT 6
US-09-514-573-1
; Sequence 1, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-1

Alignment Scores:

Pred. No.: 3.01e-43 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-514-573-1 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGAAAGGCAGCCCTGGTTGGCATTTATGGCAGAAAAGGCTCCCTTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGAGGCTCCTAAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1893

RESULT 7
US-09-252-991A-11841/c
; Sequence 11841, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11841
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11841

Alignment Scores:
Pred. No.: 1.87 Length: 1083
Score: 69.00 Matches: 28
Percent Similarity: 44.00% Conservative: 5
Best Local Similarity: 37.33% Mismatches: 30
Query Match: 18.55% Indels: 12
DB: 4 Gaps: 4

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-11841 (1-1083)

QY 9 IleAspHisProValHisLeuLysPheAsnValSerLeuGlyLysAlaAlaLeuVal--- 27
Db 1081 ATTGATCACACTGGCCACCTGGGCATCCAGCTGGGCTGGGCTTGTGCGCC 1022

QY 28 GlyIleTyrGlyArgLysGlyLeuProPro-----SerHisThrGln 41
Db 1021 GGCCTGGCGGTGCTGCCGACACCCGCCCTCGAACAGCGGACGCTGAGACGAG 962

QY 42 PheAsp-----PheValGluLeuLeuAspGly-----ArgArgLeuLeuThrGln 56
Db 961 GCCGACACTGTCGGTGTGCTATTCTGCTCGTAGTGTTCGGAAGCCGAGTGCCT 902

QY 57 GluAlaArgSerLeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 901 GCTGCTGCTGGCGTAGAGGCTGAGCCTGGGCAAGTGGCGCGGCACG 857


```

RESULT 8
US-09-252-991A-11543
; Sequence 11543, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11543
; LENGTH: 1407
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11543

Alignment Scores:
Pred. No.:          2.67          Length:      1407
Score:              69.00          Matches:      28
Percent Similarity: 44.00%          Conservative: 5
Best Local Similarity: 37.33%          Mismatches: 30
Query Match:        18.55%          Indels:      12
DB:                  4              Gaps:         4

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-11543 (1-1407)

QY      9 IleAspHisProValHisLeuLysPheAsnValSerLeuGlyLysAlaAlaLeuVal--- 27
      |||||
Db      328 ATTGATCACACTGGCCACCTGGGCATCCAGCTCGGCCTGGCGCTACTTGTGCCC 387
      |||||

QY      28 GlyIleTyrGlyArgLysGlyLeuProPro-----SerHisThrGln 41
      |||||
Db      388 GGCCTGGCGGTGCTGCCGACACCCGCCCTCGAACAGCGGCAGGCTGAGACGCAG 447
      |||||

QY      42 PheAsp-----PheValGluLeuLeuAspGly-----ArgArgLeuThrGln 56
      |||||
Db      448 GCCGACACTGTGGTGTGCTGATTCTGCTCGTAGGTGCTTCCGAAGCCGAGTGCCTTT 507
      |||||

QY      57 GluAlaArgSerLeuGluGlyThrProArgGlnSerArgGlyThr 71
      |||||
Db      508 GCTGCTGCTGGCTAGAGGCTGAGCCTGGGCAAGTGGCCGGCACG 552
      |||||

RESULT 9
US-09-252-991A-11691/c
; Sequence 11691, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11691
; LENGTH: 2304
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11691

Alignment Scores:
Pred. No.:          5.21          Length:      2304
Score:              69.00          Matches:      28
```

```

Percent Similarity: 44.00%          Conservative: 5
Best Local Similarity: 37.33%          Mismatches: 30
Query Match:        18.55%          Indels:      12
DB:                  4              Gaps:         4

US-10-029-020-14_COPY_450_520 (1-71) x US-09-252-991A-11691 (1-2304)

QY      9 IleAspHisProValHisLeuLysPheAsnValSerLeuGlyLysAlaAlaLeuVal--- 27
      |||||
Db      1977 ATTGATCACACTGGCCACCTGGGCATCCAGCTCGGCCTGGGCCTGGCGTACTTGTGCCC 1918
      |||||

QY      28 GlyIleTyrGlyArgLysGlyLeuProPro-----SerHisThrGln 41
      |||||
Db      1917 GGCCTGGCGGTGCTGCCGACACCCGCCCTCGAACAGCGGCAGGCTGAGACGCAG 1858
      |||||

QY      42 PheAsp-----PheValGluLeuLeuAspGly-----ArgArgLeuThrGln 56
      |||||
Db      1857 GCCGACACTGTGGTGTGCTATTCTGCTCGTAGGTGCTTCCGAAGCCGAGTGCCTTT 1798
      |||||

QY      57 GluAlaArgSerLeuGluGlyThrProArgGlnSerArgGlyThr 71
      |||||
Db      1797 GCTGCTGCTGGCTAGAGGCTGAGCCTGGGCAAGTGGCCGGCACG 1753
      |||||

RESULT 10
US-09-669-751-81
; Sequence 81, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Ralph J.
; TITLE OF INVENTION: Methods for Identifying Compounds for
; TITLE OF INVENTION: Motion Sickness, Vertigo and Other Disorders Related to
; TITLE OF INVENTION: Balance and the Perception of Gravity
; FILE REFERENCE: P-NI 3864
; CURRENT APPLICATION NUMBER: US/09/669,751
; CURRENT FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US 60/168,579
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 261
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Drosophila
US-09-669-751-81

Alignment Scores:
Pred. No.:          1.14          Length:      523
Score:              67.50          Matches:      19
Percent Similarity: 56.52%          Conservative: 7
Best Local Similarity: 41.30%          Mismatches: 17
Query Match:        18.15%          Indels:       3
DB:                  4              Gaps:         2

US-10-029-020-14_COPY_450_520 (1-71) x US-09-669-751-81 (1-523)

QY      23 LysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGlnPhe 42
      |||||
Db      340 AAGGTGTGTGGTGGTGTATTATACCAAGGATGGCGATAAGCCATCGAAACCCGCCA 399
      |||||

QY      43 AspPheValGluLeuLeuAsp-----GlyArgArgLeuThrGln---GluAlaArg 59
      |||||
Db      400 AATGCAGTGACCTTGGATGATGCTGTGGTGGCAAGCTGTTGACCCCTGATCCGTGAACGT 459
      |||||

QY      60 SerLeuGluGlyThrPro 65
      |||||
Db      460 GGAATGGACGGCACTCCC 477
      |||||

RESULT 11
US-09-252-991A-14082/c
; Sequence 14082, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
```


; OTHER INFORMATION: Incyte ID No. 6682888 196545.7
; NAME/KEY: unsure
; LOCATION: 3743-3768
; OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-120

Alignment Scores:
Pred. No.: 28 Length: 3835
Score: 66.00 Matches: 19
Percent Similarity: 40.00% Conservative: 11
Best Local Similarity: 25.33% Mismatches: 25
Query Match: 17.74% Indels: 20
DB: 4 Gaps: 2

US-10-029-020-14_COPY_450_520 (1-71) x US-09-566-921-120 (1-3835)

QY 11 HisProValHisLeu----- 15
Db 2940 CACCCTACTCATCTTACACCTGCTCTAGACAAAGTTGTGAATGACTTAAAAAATACTT 2881
QY 16 ---LysPheAsnValSerLeuGlyLysAlaLeuValGlyIleTyrGlyArgLysGly 34
Db 2880 TTTCAATTTCCTTTGGCATAACCACTGCTGGACTGCTGGGTTCACATGGA----- 2830
QY 35 LeuProProSerHisThrGlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeu 54
Db 2829 ---CCTACGATTTCCCTCCACATGCCCTTTATTGATCTGGGTCTGGTAAGTTGATGCTA 2773
QY 55 ThrGlnGluAlaArgSerLeuGluGlyThrProArgGlnSerArg 69
Db 2772 ATGGCAGAAGCAATAGAAAGTCCAAGGCACTAAAAATGTCCAGCAAG 2728

RESULT 15
US-09-584-568C-3
; Sequence 3, Application US/09584568C
; Patent No. 6500657
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria, Alexandra et al.
; TITLE OF INVENTION: 33167, A NOVEL HUMAN HYDROLASE AND USES THEREFOR
; FILE REFERENCE: MNI-140
; CURRENT APPLICATION NUMBER: US/09/584,568C
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 60/193,954
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 945
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(945)
US-09-584-568C-3

Alignment Scores:
Pred. No.: 5.81 Length: 945
Score: 65.00 Matches: 19
Percent Similarity: 53.45% Conservative: 12
Best Local Similarity: 32.76% Mismatches: 23
Query Match: 17.47% Indels: 4
DB: 4 Gaps: 1

US-10-029-020-14_COPY_450_520 (1-71) x US-09-584-568C-3 (1-945)

QY 12 ProValHisLeuLysPheAsnValSerLeuGlyLysAlaLeuValGlyIleTyrGly 31
Db 154 CCGCTTCGCTTCTCCTACAGGCTTCTGGACGGGGAGGAGCCCTCCCGCCGTCGTCTTT 213
QY 32 ArgLysGlyLeuProProSerHisThrGlnPheAspPheValGluLeu----- 48
Db 214 TTGCACGGGCTCTTCGGCAGCAAACTAACTCAACTCCATCGCAAGATCTTGGCCCGAG 273

QY 49 ---AspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeuGluGlyThrPro 65
Db 274 CAGACAGGCCGTAGGTGCTGACGGTGGATGCTCGTAACCAACGGTGACAGCCCC 327

Search completed: August 14, 2004, 19:39:45
Job time : 37.0053 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 02:35:12 ; Search time 534 Seconds
(without alignments)
3182.171 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76
Perfect score: 2132
Sequence: 1 MDVKKRKPYSLTRRDAER.....EITEDTASSWPVPTDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/MITRA020/runat_06082004_114100_196/app_query.fasta.1.583
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MITRA020 @CGN_1_1_470 @runat_06082004_114100_196 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	2127	99.8	8354	6	ABS52100
2	2122	99.5	3111	2	AAV19251
3	2122	99.5	3111	2	AAX87705
4	2122	99.5	3111	3	AAD00791
5	2122	99.5	3111	4	AAS18526
6	2082	97.7	8438	6	ABN85378
7	2066	96.9	8645	6	ABS78652
8	1285	60.3	2387	2	AAV19252

9	975	45.7	1534	9	ADE07179	Ade07179	Novel	cod
10	975	45.7	8645	6	ABQ82344	Abq82344	Human	NOV
11	975	45.7	8675	6	ABQ82343	Abq82343	Human	NOV
12	937	43.9	9695	7	ACC72052	Acc72052	BCU0205B	
13	828	38.8	9729	5	AAS14089	Aas14089	Human	FCT
14	828	38.8	9729	9	ADB32028	Adb32028	Human	FCT
15	828	38.8	9826	5	AAS14085	Aas14085	Human	FCT
16	828	38.8	9826	9	ADB32023	Adb32023	Human	FCT
17	741.5	34.8	13202	4	AAK51828	Aak51828	Human	pol
18	733.5	34.4	12879	6	ABK92230	Abk92230	Prostate	
19	641.5	30.1	8473	6	ABQ82345	Abq82345	Human	NOV
20	641.5	30.1	8487	6	ABQ82346	Abq82346	Human	NOV
21	527	24.7	3052	6	ABQ75971	Abq75971	Human	pai
22	522	24.5	1755	4	AAF23418	Aaf23418	Human	SEC
23	521.5	24.5	978	9	ADE08934	Ada53142	Human	DNA
24	508	23.8	1727	7	ADA53142	Ada53142	Human	cod
25	508	23.8	9058	7	ACC72051	Acc72051	BCU0205A	
26	434	20.4	1429	8	ACD40264	Acd40264	Human	bre
27	434	20.4	1430	5	AAS14084	Aas14084	Human	FCT
28	434	20.4	1430	9	ADB32021	Adb32021	Human	FCT
29	434	20.4	1431	4	AAF27861	Aaf27861	Human	NOV
30	434	20.4	1743	5	AAS01213	Aas01213	DNA	encod
31	404	18.9	2676	5	AAS67421	Aas67421	DNA	encod
32	332.5	15.6	452	6	ABS72992	Abs72992	Human	gen
33	279	13.1	328	2	AAV87144	Aav87144	EST	clone
34	279	13.1	1692	7	ACD05836	Acd05836	Novel	hum
35	277	13.0	643	5	AAS67419	Aas67419	DNA	encod
36	167.5	7.9	784	5	AAS69256	Aas69256	DNA	encod
37	141	6.6	9222	7	AAL51171	Aal61171	Actinosyn	
38	141	6.6	82746	7	AAL61224	Aal61224	Actinosyn	
39	139	6.5	6015	4	ABL06863	Abl06863	Drosophil	
40	139	6.5	6815	4	ABL06862	Abl06862	Drosophil	
41	138.5	6.5	24533	4	AAS27689	Aas27689	DNA	encod
42	138.5	6.5	24533	7	ABT17017	Abt17017	Human	sec
43	138.5	6.5	24533	7	ABZ68157	Abz68157	Human	sec
44	138.5	6.5	24533	7	ABZ74637	Abz74637	Secreted	
45	138.5	6.5	24533	9	ADB94492	Adb94492	Novel	hum

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
XX
AC ABS52100;
XX
DT 05-NOV-2002 (first entry)
XX
DE Human TEN-M4-like gene.
XX

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
cell signal processing; metabolic pathway modulation; metabolic disorder;
obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
memory defect; infertility; congenital heart defect; hair growth;
pigmentation disorder; endocrine disorder; respiratory disease; health;
gastro-intestinal disease; reproductive; neurological disease;
bone marrow transplantation; endocrine disease; allergy; inflammation;
neurological disorder; urinary system disorder; age-related disorder;
neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
adipocyte complement-related C1q tumour necrosis factor; out at first;
beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
type 1a membrane sushi domain containing; butyrophilin;
single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers
variation replace(117,G)
/*tag= a

```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US050331.
XX
PR 19-DEC-2000; 2000US-0256704P.
PR 20-DEC-2000; 2000US-0257314P.
PR 02-MAY-2001; 2001US-0288153P.
PR 29-MAY-2001; 2001US-0294075P.
PR 24-JUL-2001; 2001US-0307506P.
PR 10-AUG-2001; 2001US-0311590P.
PR 10-AUG-2001; 2001US-0311613P.
PR 29-AUG-2001; 2001US-0315617P.
PR 14-SEP-2001; 2001US-0322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.
XX
PS Claim 8; Page 50-52; 318pp; English.
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders. The
CC present nucleic acid sequence represents a NOVX gene. This sequence
CC encodes a NOVX protein of the invention
XX
SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 6.54e-127 Length: 8354
Score: 2127.00 Matches: 397
Percent Similarity: 99.25% Conservative: 0
Best Local Similarity: 99.25% Mismatches: 3
Query Match: 99.77% Indels: 0
DB: 6 Gaps: 0
XX
SEQ14-X-AT-28-64-76 (1-400) x ABS52100 (1-8354)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
Db 35 ATGGACGTGAAGGAGAGGAGACCTTACCGCTCGCTGACCCGGCGCCGACGCCGAGCGC 94
```

```
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 95 CGCTACACAGCTCGTCCGGGACAGCGAGGAGGCAAGCCCCCGCAGAAATCGTACAGC 154
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 155 TCCAGCGAGACCTGAAGGCTACGACCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC 214
QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 215 AAGGACATTGTCCCGCAGAGGCCGAGGAATTCTGCCGACAGGTGCCAACTTCACCCCTG 274
QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 275 CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCTGTACCGGACAGACATT 334
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 335 GGCTGCCCAATGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 394
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 395 ACGTGCTGTCCCTGAGCACCCCGTCGCTGTGTGGGCGGAGCACACGGTCAGGGCGC 454
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 455 AGCTCTGCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 514
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 515 AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCCACGGCGGCTCCGGACGCCCGC 574
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 575 CCGCCGCTCTCGCACGCCCCACACCCCAACAGCACACCGCGGCTCCATTAACTCCCTG 634
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 635 AACCGGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCACGACCGACCTCGCTC 694
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 695 TCCGGAGAGCCCTCGCGGCGCGCCAGAGCCTGCCACCGCCAGGAGAACTGGCTG 754
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 755 CTCAACAGCAACATCCCCCTGGAGACCAGGAACCTAGGCAAGCAGCCATTCTTAGGGACA 814
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 815 TTGCAGGACAACTCATTTGAGATGGACATTCTCGCGCCTCCCGCCATGATGGGCTTAC 874
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 875 AGTGACGGGCACTTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGACCCACATCA 934
QY 301 proGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 935 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGCCCGCTGCCCGC 994
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 995 AGCACCTTCGCCCCGGCGCCTTTAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1054
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1055 TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTTGGTTCATCTGTCGTCATACTTT 1114
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1115 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTCGAGCCGATGGAGGGCAGATGTAT 1174
```


QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Dd 1175 GAGATCACGGAGGACACAGCCAGCAGTTGGCCTGCGCAACCGACGCTCTCCCTATACCCC 1234

RESULT 2
AAV19251
ID AAV19251 standard; cDNA; 3111 BP.
XX
AC AAV19251;
XX
XX
DT 17-AUG-1998 (first entry)
XX
DE Human gamma-hereregulin cDNA.
XX
KW Gamma-hereregulin; gamma-HRG; human; autocrine growth factor;
KW breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
KW cell proliferation; cell differentiation; cell survival;
KW neurological disorder; muscular disorder; ss.
XX
OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 534..2645
FT /*tag= a
XX
PN WO9802541-A1.
XX
PD 22-JAN-1998.
XX
XX 08-JUL-1997; 97WO-US011841.
PF
XX
PR 12-JUL-1996; 96US-0021640P.
XX
XX (GBTH) GENENTECH INC.
PA
XX Schaefer GM, Sliwkowski M;
PI
XX
XX WPI; 1998-110589/10.
DR
DR P-PSDB; AAW44817.
XX

DNA encoding gamma-hereregulin - used to activate ErbB receptor and to enhance proliferation, differentiation or survival of a cell.
PT
PT
XX
PS Claim 21; Fig 1A-C; 81pp; English.

XX This nucleic acid molecule codes for human gamma-hereregulin (gamma-HRG) (see AAW44817), a novel member of the heregulin superfamily, that has a unique N-terminal domain not present in previously identified heregulins. Gamma-HRG is secreted by human breast cancer MDA-MB-175 cells. It leads to the formation of a constitutive active receptor complex and stimulates the growth of these cells in an autocrine manner. The nucleic acid was isolated by screening a MDA-MB-175 with a probe corresponding to the EGF-like domain and part of the N-terminal sequences of HRG-beta-3. It can be used for the recombinant production of gamma-HRG, or for in vivo or ex vivo gene therapy. A claimed nucleic acid, which is complementary to the nucleic acid sequence encoding the N-terminal domain of gamma-HRG, is able to reduce production of gamma-HRG by MDA-MB-175 cells. The invention provides claimed methods for activating an ErbB receptor and for enhancing proliferation, differentiation or survival of a cell by contacting the cell (preferably a glial or muscle cell) with gamma-HRG

SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.42e-127 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 2 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AAV19251 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Dd 334 ATGGACGTGAAGGAGAGAACGCTTACCGCTCGCTGACCCGGCGCCGACGCGAGCGC 393
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Dd 394 CGCTACACCAGCTCGTCCGCGGACAGCGAGGAGGCAAGCCCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Dd 454 TCCAGCGAGACCCCTGAAGGCCTACGACCAGGAGCGCCGCTAGCCTATGGCAGCGCGTC 513
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Dd 514 AAGGACATTGTGCCGAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCCACTTCACCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Dd 574 CGGAGCTGGGGCTGGAAGAAGTAACGCCCTCACGGGACCTGTACCGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Dd 634 GGCCTCCCCACTGCGCTACTCCATGGGGCTGGCTGTATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Dd 694 ACGGTGCTGTCCCTGAGCACCCCGTGCCTGTCTGGGGCCGGAGCACACGCTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Dd 754 AGCTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Dd 814 AACACTGAGACTGATCATCCGGGGCGGCTCGAGAACCAACGCGCGGCTCCGACGCCGCGC 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Dd 874 CCGCCGCTCTCGCACGCCCCACACCCCCAACCCAGCACCAACGCGCGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Dd 934 AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCGCCACGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Dd 994 TCCGGAGAGCCCCCTGCGGGCGGCCCGAGAGCTGCCACGCCCGCAGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Dd 1054 CTCAACAGCAACATCCCCCTGGAGACCCAGAAACCTAGGCAAGCAGCCATCTCTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Dd 1114 TTGCAGGACAACTCATTTGAGATGGACATTTCTGGCGCCTCCCGCCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Dd 1174 AGTGACGGGCACTTCTCTTCAAGCCTTGAGGACACTTCCCGCTCTTCTGACCAACATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Dd 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGCCAGCCCTGCCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Dd 1294 AGCACCTTCGCCCGCGCGGCTTTAACCTCAAGAAGCCCTCCAAGTACTGTAAGTGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Dd 1354 TGCGCAGCCCTGAGCGCCATCGTCTATCTCAGCCACTCTGGTCTCTCTGCTGCTACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1414 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473
Qy 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 3
AAX87705
ID AAX87705 standard; cDNA; 3111 BP.
XX
AC AAX87705;
XX
DT 26-OCT-1999 (first entry)
XX
DE Gamma-herregulin cDNA.
XX
KW Gamma-herregulin; HRG-alpha; human; ligand; HER2; HER3; HER4; receptor;
KW lung surfactant; respiratory distress syndrome; emphysema;
KW epithelial growth factor; therapy; ss.

XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 334..2640
FT /*tag= a

PN WO9939729-A2.

XX

PD 12-AUG-1999.

XX

PF 03-FEB-1999; 99WO-US002390.

XX

PR 04-FEB-1998; 98US-00020598.

XX

PA (GETH) GENENTECH INC.

PA (IOWA) UNIV IOWA RES FOUND.

XX Sliwkowski M, Kern JA;

XX WPI; 1999-494213/41.

DR P-PSDB; AAY06639.

XX

PT Heregulin ligands can be used to induce epithelial cell growth, and to

PT promote repair and healing of tissue damage or injury.

XX Disclosure; Page 111-116; 120pp; English.

CC This is the nucleotide sequence of cDNA coding for gamma-herregulin (gamma
CC -HRG, see AAY06639). The invention provides HRG ligands, including gamma-
CC HRG, that have affinity for and stimulate HER2, HER3 and/or HER4
CC receptors in autophosphorylation. A new method of treating respiratory
CC distress syndrome in humans uses HER2, HER3 and/or HER4 receptor ligands
CC as epithelial growth factors. A novel method of inducing epithelial cell
CC growth and/or proliferation comprises contacting a normal epithelial cell
CC which expresses HER2, HER3 and/or HER4 receptors with an isolated ligand
CC which activates HER2, HER3, HER4 receptors or their combination. Also
CC claimed are methods of increasing lung surfactant protein A, or of
CC treating chronic obstructive pulmonary disease, respiratory distress or
CC emphysema, by administering an effective amount of an isolated HER ligand
CC to a patient

SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.42e-127 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 2 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AAX87705 (1-3111)

QY 1 MetAspVallysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
Db ATGGACGTGAAGAGAGAGAACCTTACCGCTTCGTGACCCGGGCGCGGACCGGAGCGC 393
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db CGCTACACACAGCTCGTCCGCGGACAGCGAGAGGGCAAAGCCCCGAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db TCCAGCGAGACCTGAAGGCCTACGACCGAGGCGCCCGCTAGCCTATGGCAGCGCGCTC 513
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db AAGGACATTGTGCCCGCAGGAGCGCGAGGAATTCTGCCGCACAGGTGCCCACTTCACCCTG 573
QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db CGGAGCTGGGGCTGGAAGAAAGTAACGCCCCCTCACGGGACCTGTACCGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db GGCCTCCCCACTGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db ACGGTGCTGTCCCTGAGCACCCCGTGCCTCTGTGGGGCCGGAGCACACGGTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db AGCTCCTGCCTGTCCAGCCGGGCAATTCCAAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db AACACTGAGACTGATCATCCGGGGCGCTGCAGAACACCGCGGCTCCGGACGCGCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db CCGCGCTCTCGCACGCCACACCCCCAACCCAGCACACCGCGGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCACCGGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db TCCGAGAGCCCCCTGCCGGCGCGCCAGGAGCTGCCCCACGCCACGAGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db CTCAACAGCAACATCCCTCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGCGCTCCCGCCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db AGTGACGGGCACCTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGCCCTGCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db AGCACCTTCGCCCGCGCGCTTTAACCTCAAGAAGCCCTCCAAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db TCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTGCTGGCATACTTT 1413

[illegible]

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
DQ |||||||
DB 1294 AGCACCTTCGCCCGCGGCTTTAACCTCAAGAAGCCCTCCAAGTACTGTAAGTGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
DB |||||||
DB 1354 TGGCAGCCCTGAGCGCATCGTCATCTCAGCCACTCTGGTCATCTCTGCTGSCATACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
DB |||||||
DB 1414 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGACATGTAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
DB |||||||
DB 1474 GAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCCTCCCTATACCCC 1533

RESULT 5
AAS18526
ID AAS18526 standard; cDNA; 3111 BP.
XX
AC AAS18526;
XX
DT 26-FEB-2002 (first entry)
XX
DE DNA encoding human heregulin, gamma-HRG.

XX Human; heregulin; antiasthmatic; antiinflammatory; vulnery; antiulcer;
FH epithelial cell growth; HER2; HER3; HER4; HRG; lung cell; asthma;
FT lung surfactant protein A; respiratory distress; emphysema; pneumonitis;
FT chronic obstructive pulmonary disease; COPD; chronic bronchitis; ulcer;
XX neonatal pulmonary disease; neonatal respiratory distress syndrome;
XX meconium aspiration syndrome; congenital diaphragmatic hernia; lesion;
XX acute lung injury; cystic fibrosis; surgical wound; resection;
XX growth factor; smoke inhalation; gamma-HRG; ss.
XX Homo sapiens.

FT Key Location/Qualifiers
FH CDS 334..2640
FT /*tag= a
FT /product= "Human heregulin, gamma-HRG"
XX US2001023241-A1.
XX
XX 20-SEP-2001.
XX
XX 02-FEB-2001; 2001US-00773517.
XX
XX 04-FEB-1998; 98US-0073866P.
XX 02-FEB-1999; 99US-00243198.
XX (SLIW/) SLIWKOWSKI M X.
XX (KERN/) KERN J A.
XX
XX Sliwkowski MX, Kern JA;
XX WPI; 2001-595807/67.
XX P-PSDB; AAU09891.
XX
XX Inducing epithelial cell growth and/or proliferation, useful in the
XX treatment of respiratory disease, comprises use of heregulin ligand as
XX growth factor.
XX
XX Disclosure; Fig 7; 68pp; English.

XX The invention relates to inducing epithelial cell growth and/or
XX proliferation comprising contacting a normal epithelial cell, which
XX expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4,
XX HER3/HER4, HER3 or HER4) receptors with an isolated ligand such as
XX heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The
XX method is used for inducing epithelial cell (preferably lung cell) growth
XX and/or proliferation, for increasing lung surfactant protein A and for
XX treating respiratory distress or emphysema, for treating chronic

CC obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma,
CC neonatal pulmonary diseases including neonatal respiratory distress
CC syndrome, meconium aspiration syndrome, chronic lung disease of the
CC neonate, congenital diaphragmatic hernia and acute lung injuries
CC including smoke or chemical inhalation, pneumonitis due to aspiration,
CC radiation, near drowning, cystic fibrosis and other epithelial cell
CC trauma diseases including injuries associated with surgical wounds and
CC resections, ulcers, lesions and tissue tears as normal epithelial cell
CC growth factors. For treating infants/neonates with respiratory distress
CC as well as youth and adult with poor lung function due to lung injury or
CC damage. The HRG binds with varying and very high affinity to the HER2,
CC HER3 and/or HER4 receptors. The method stimulates growth and
CC proliferation of the epithelial cells, repairing and re-establishing the
CC cellular barriers of organs and allowing the affected tissue to develop
CC normal physiological functions more quickly. Hence the method improves
CC oxygenation and speeds of the development of a barrier to infection while
CC treating lung cells that are damaged by inhalation of smoke resulting in
CC emphysema. It also facilitates regeneration of epithelial cells. The
CC present sequence represents the coding sequence of human heregulin gamma-
CC HRG
XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.42e-127 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 4 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x AAS18526 (1-3111)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
DB 334 ATGGACGTGAAGGAGAGAGCCCTTACCGCTCGCTGACCGCGCGCGAGCGC 393
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 394 CGCTACACCAGCTCGTCCGGGACAGCGAGGAGGGGAAAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 454 TCCAGCGAGACCTTGAGGCCTACGACCAGGACGCGCGCTAGCCTATGGCAGCGCGTC 513
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly**AsnPheThrLeu 80
DB 514 AAGGACATTGTGCCGAGGAGCGCGAGGAATTCGCGCACAGGTGCCAACTTCACCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
DB 574 CGGAGCTGGGGCTGGAGAAAGTAACGCCCTCACGGACCCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
DB 634 GGCCTCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
DB 694 ACGTGCTGTCCCTGAGCACCCCGTGTGCTGTGGGCGCGGAGCACACGGTCAGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB 754 AGCTCCTGCTGTCCAGCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
DB 814 AACACTGAGACTGATCATCCGGCGGCTGCAGAACACACGCGCGGCTCCGAGCGCGCGC 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
DB 874 CCGCGCTCTCGCACGCGCCACACACCCCAACACGACACCGCGCGCTCCATTAACTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCGCCACCGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGCCCCCTGCGGGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGGCGCTCCCGCCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACCTTCCTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCACATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCCTCCGCCCGACCCCTGCCCCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGGCCCTTTAAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGCAGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCCTGTGGCATACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCGATGGAGGGGCAGATGTAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGTCTCCCTATACCCC 1533

RESULT 6
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
XX
AC ABN85378;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human NOV1, TEN-M4 like protein, coding sequence.
XX
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 4. .8395
FT /*tag= a
FT /trans_except= (pos: 1138. .1147,aa:Met)
FT /product= "NOV1 protein"

PN WO200255704-A2.
XX
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-US000554.

XX 09-JAN-2001; 2001US-0260417P.
PR 10-JAN-2001; 2001US-0260831P.
PR 28-FEB-2001; 2001US-0272338P.
PR 09-MAR-2001; 2001US-0274876P.
PR 18-APR-2001; 2001US-0284704P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
PI Gunther E, Stone DJ;
XX
DR WPI; 2002-590674/63.
DR P-PSDB; ABB98401.
XX
PT NOVX polypeptides and encoding polynucleotides, useful for preventing or
PT treating NOVX-associated disorders e.g. cancer, inflammation, or
PT Alzheimer's disease, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 9; Page 8-9; 358pp; English.
XX
CC The present sequence is a coding sequence for a NOV protein. The NOV
CC proteins and coding sequences are useful for treating or preventing NOV-
CC associated disorders or in the manufacture of a medicament for treating
CC the disorders, such as cancer, heart disease, inflammation, autoimmune
CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
CC and other wasting disorders associated with chronic diseases. NOV1 is a
CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX
SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.17e-124 Length: 8438
Score: 2082.00 Matches: 395
Percent Similarity: 98.01% Conservative: 0
Best Local Similarity: 98.01% Mismatches: 5
Query Match: 97.65% Indels: 3
DB: 6 Gaps: 1

SEQ14-X-AT-28-64-76 (1-400) x ABN85378 (1-8438)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 4 ATGGACGTGAAGGAGAGGAGCCTTACCGCTCGTGACCCGCGCCGACGCGAGCGC 63
QY 21 ArgTyrThrSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 64 CGCTACACCAAGCTCGTCCGCGGACAGCGAGGAGGCAAAAGCCCCGAGAAATCGTACAGC 123
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 124 TCCAGCGAGACCCTGAAGGCTCTACGACCAGGAGCGCCCGCTAGCCTATGGCAGCGCGTC 183
QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 184 AAGGACATTGTGCGCGAGGAGCGCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG 243
QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 244 CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCTGTACCGACAGACATT 303
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 304 GGCCTCCCCACTGCGGTACTCTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 363
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140

Db	364	ACGGTGTCTGCCCTGAGACACCCCGTGGCTCTGTGGGCGGAGCACACGGTCAGGGCGC	423	KW	renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
Qy	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160	KW	neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
Db	424	AGTCTCTGCCTGTCCAGCGGGCCAATTCCAAATCTCACACTCACCGACACCGAGCATGAA	483	KW	reproductive disorder; infertility; autoimmune disorder; gout; allergy;
Qy	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180	KW	inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
Db	484	AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACACGCGGGCTCCGGACGCGCGCG	543	KW	autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
Qy	181	ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu	200	KW	diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
Db	544	CGCGGCTCTCGACGCGCCACACCCCAACACGACACGCGGCTCCATTAACTCCCTG	603	KW	multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
Qy	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220	KW	rheumatoid arthritis.
Db	604	AACCGGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGGCCCCACGGACCACTCGCTC	663	XX	Homo sapiens.
Qy	221	SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240	PN	WO200272830-A2.
Db	664	TCCGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG	723	XX	19-SEP-2002.
Qy	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260	PF	08-FEB-2002; 2002WO-US003715.
Db	724	CTCAACAGCAACATCCCTCGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGACA	783	XX	09-FEB-2001; 2001US-0268111P.
Qy	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280	PR	23-FEB-2001; 2001US-0271175P.
Db	784	TTGCAGGACAACCTCATTGATGGACATTCTCGGCGCCTCCCGCCATGATGGGGCTTAC	843	PR	08-MAR-2001; 2001US-0274503P.
Qy	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300	XX	09-MAR-2001; 2001US-0274552P.
Db	844	AGTGACGGGCACTTCCTCTTAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCACTCA	903	PA	(INCY-) INCYTE GENOMICS INC.
Qy	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320	XX	Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
Db	904	CCAGGTACCCACTGACGTCACGACAGTGTACTCTCTCCGCCCGACCCCTGCCCGCG	963	PI	Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
Qy	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340	PI	Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
Db	964	AGCACTTCGCCTGGCCGGCCCTTAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAG	1023	XX	WPI; 2002-723356/78.
Qy	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360	DR	P-PSDB; ABG97359.
Db	1024	TGCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCCTGTGGCATACTTT	1083	XX	New human proteins associated with cell growth, differentiation and
Qy	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet---	379	PT	death, useful for diagnosing, treating or preventing autoimmune or
Db	1084	GTGGCCATGCACCTGTTTGGCCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGACG	1143	PT	inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
Qy	380	----TyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe	398	XX	atherosclerosis or hepatitis.
Db	1144	GATTTATGAGATCACGGAGGACACAGCCAGAGTTGGCCTGTGCCAACCGACGCTCCCT	1203	XX	Claim 5; Page 175-178; 181pp; English.
Qy	398	uTyrPro 400		XX	The invention relates to an isolated polypeptide comprising CGDD1-12
Db	1204	ATACCCC 1210		CC	(cell growth, differentiation and death), a naturally occurring amino
RESULT 7				CC	acid sequence at least 90% identical to CGDD, a biologically active
ABS78652				CC	fragment or an immunogenic fragment. Also included are the
XX				CC	polynucleotides encoding CGDD1-12, a recombinant polynucleotide
AC				CC	comprising a promoter sequence operably linked to the CGDD
XX				CC	polynucleotides, a cell transformed with the recombinant polynucleotide,
XX				CC	a transgenic organism comprising the recombinant polynucleotide, an anti-
XX				CC	CGDD antibody, screening for compounds which bind to/modulate or are
DT				CC	ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
XX				CC	CGDD polynucleotide microarray. The polypeptides, polynucleotides,
DE				CC	agonists and antagonists are useful for diagnosing, treating or
XX				CC	preventing disorders associated with aberrant expression of CGDD,
KW				CC	particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
KW				CC	cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
KW				CC	vera, psoriasis, primary thrombocytopaenia or cancer), developmental
KW				CC	disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
KW				CC	neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
KW				CC	epilepsy), reproductive disorders (e.g. infertility or a disruption in
KW				CC	the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
KW				CC	acquired immunodeficiency syndrome) allergies, asthma, autoimmune
KW				CC	thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
KW				CC	glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
KW				CC	osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
KW				CC	uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
KW				CC	infections. They are also useful in the assessment of the effects of
KW				CC	exogenous compounds on the expression of nucleic acid and amino acid
KW				CC	sequences of proteins associated with CGDD. The present sequence encodes
KW				CC	a CGDD protein
DE				XX	Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;
XX				SQ	
Alignment Scores:					
Pred. No.:					5.67e-123
Score:					2066.00
					Length: 8645
					Matches: 390

Percent Similarity: 97.50%				Conservative: 0			
Best Local Similarity: 97.50%				Mismatches: 6			
Query Match: 96.90%				Indels: 4			
DB: 6				Gaps: 2			
SEQ14-X-AT-28-64-76 (1-400) x ABS78652 (1-8645)							
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg	20				
Db	117	ATGGACGTGAAGAGAGAGAACCTTACCGCTCGCTGACCCGGCGCGCGAGCGCGC	176				
QY	21	ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer	40				
Db	177	CGCTACACACAGCTCGTCCGGGACAGCGAGGAGGGCAAAGCCCCGAGAAATCGTACAGC	236				
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60				
Db	237	TCCAGCGAGACCTGAAGGCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCGCGTC	296				
QY	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80				
Db	297	AAGGACATTGTGCCGAGGCGGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG	356				
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100				
Db	357	CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT	416				
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120				
Db	417	GGCCTCCCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	476				
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140				
Db	477	ACGGTGCTGTCCCTGAGCACCCCGTGCCTGTGTGGGCGCGGAGCACACGGTCAGGGCGC	536				
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160				
Db	537	AGCTCCTGCCTGTCCAGCGCGGCCAAATCCAAATCTCACTACCGACACCGAGCATGAA	596				
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180				
Db	597	AACACTGAGACT-----CCGGGCGGCTGCAGAACCCACGCGGGCTCCGGACGCGCGCG	650				
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200				
Db	651	CGCGCGCTCTCGCACGCCCAACCCCCAACCCAGCACCAAGCGCGGCTCCATTAACCTCCCTG	710				
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220				
Db	711	AACCGGGGCAACTTCACGCGAGGAGCAACCCCGCGCGCCCGCCACCGACCACTCGCTC	770				
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240				
Db	771	TCCGAGAGCCCCCTGCGGCGCGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG	830				
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260				
Db	831	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTCTAGGGACA	890				
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280				
Db	891	TTGCAGGACAACCTCATTTAGATGGACATTCTCGCGGCTCCCGCCATGATGGGGCTTAC	950				
QY	281	SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer	300				
Db	951	AGTGACGGGCACTTCTCTTCAAGCCTGGAGGACCTCCCCGCTCTTCTGTGACCACATCA	1010				
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg	320				
Db	1011	CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGGACCCCTGCCCGC	1070				
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340				

Db	1071	AGCACCTTCGCCCGCGCGCCTTTAAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG	1130				
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360				
Db	1131	TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTGCTGGCATACTTT	1190				
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380				
Db	1191	GTGGGTAAAGCACTCTTC-----AACTGGCACTGCAGCCGATGGAGGGCAGATGTAT	1244				
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400				
Db	1245	GAGATCACGGAGGACACAGCCAGCAGTTGGCTGTGTGCCAACCGACGCTCTCCCTATACCCC	1304				
RESULT 8							
AAV19252							
ID	AAV19252	standard; cDNA; 2387 BP.					
XX	AAV19252;						
AC	AAV19252;						
DT	17-AUG-1998	(first entry)					
XX							
DE		Human gamma-hergulin cDNA clone 20.					
XX							
KW		Gamma-hergulin; gamma-HRG; human; autocrine growth factor;					
KW		breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;					
KW		cell proliferation; cell differentiation; cell survival;					
KW		neurological disorder; muscular disorder; ss.					
OS		Homo sapiens.					
XX							
FH	Key	Location/Qualifiers					
CDS		307..1890					
FT		/*tag= a					
FT	misc_feature	1186..1263					
FT		/*tag= b					
FT		/note= "insert DNA"					
XX							
PN	WO9802541-A1.						
XX							
PD	22-JAN-1998.						
XX							
PF	08-JUL-1997;	97WO-US011841.					
XX							
PR	12-JUL-1996;	96US-0021640P.					
XX							
PA	(GETH) GENENTECH INC.						
XX							
PI	Schaefer GM, Sliwowski M;						
XX							
DR	WPI; 1998-110589/10.						
DR	P-PSDB; AAW44818.						
XX							
PT	DNA encoding gamma-hergulin - used to activate ErbB receptor and to						
XX	enhance proliferation, differentiation or survival of a cell.						
PS	Disclosure; Page 56-57; 81pp; English.						
XX							
CC	Clone 20 encodes an N-terminally truncated isoform (see AAW44818) of						
CC	human gamma-hergulin (gamma-HRG) (see also AAW44817), a novel member of						
CC	the hergulin superfamily. It was isolated from a MDA-MB-175 human breast						
CC	cancer cell library after screening with an EGF-like domain. Compared to						
CC	gamma-HRG cDNA (see AAV19251), clone 20 contains a 78 base insert between						
CC	codons 560 and 561, and lacks the 5' end. The truncated gamma-HRG was						
CC	insoluble when expressed as a thioetheroxin fusion protein in Escherichia						
CC	coli cells. Gamma-HRG nucleic acid and polypeptide are claimed, together						
CC	with their uses e.g. for enhancing the proliferation, differentiation or						
XX	survival of glial or muscle cells						
SQ	Sequence 2387 BP; 641 A; 640 C; 550 G; 556 T; 0 U; 0 Other;						
Alignment Scores:		2.04e-73	Length:	2387			
Pred. No.:							

Score:	1285.00	Matches:	235
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	60.27%	Indels:	0
DB:	2	Gaps:	0
SEQ14-X-AT-28-64-76 (1-400) x AAV19252 (1-2387)			
QY	166	HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeuSerHis	185
Db	1	CATCGGGCGGCTGCAGAACACACGCGGGCTCCGGAGCGCCGCCCGCTCTCGCAC	60
QY	186	AlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe	205
Db	61	GCCACACCCCAACCCAGCACCGGGCTCCATTAACTCCCTGAACCGGGCACTTC	120
QY	206	ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro	225
Db	121	ACGCGAGGAGCAACCCCGCGGCCCCCACGGACCTCGCTCTCCGAGAGCCCT	180
QY	226	AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle	245
Db	181	GCGCGGGCGCCAGGAGCTGCCACGCCAGGAGAACTGGTCTCAACAGCAACATC	240
QY	246	ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu	265
Db	241	CCCTGGAGACCAAGAACCTAGGCAAGCAGCCATTCTAGGACATTGCAGGACAACCTC	300
QY	266	IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe	285
Db	301	ATTGAGATGGACATTCTCGCGGCTCCCGCCATGATGGGGCTTACAGTGACGGGCACTTC	360
QY	286	LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu	305
Db	361	CTCTTCAAGCCTGGAGGACCTCCCGCTCTTGTGACCACATCACCAGGGTACCCACTG	420
QY	306	ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg	325
Db	421	ACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGCAGCACCTTCGCCCGG	480
QY	326	ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer	345
Db	481	CCGGCCTTTAACCTCAAGAGCCCTCCAAGTACTGTAACTGGAGTGGCAGCCCTGAGC	540
QY	346	AlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPheValAlaMetHisLeu	365
Db	541	GCCATCGTCATCTCAGCCACTCTGGTCACTCTGTGTCATCTTGTGGCCATGCACCTG	600
QY	366	PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp	385
Db	601	TTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTATGAGATCAGGAGGAC	660
QY	386	ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	661	ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC	705
RESULT 9			
ID	ADE07179	standard; DNA; 1534 BP.	
XX	AC	ADE07179;	
XX	XX	29-JAN-2004 (first entry)	
XX	DE	Novel coding sequence (useful for identifying genetic disorders) #245.	
XX	KW	novel gene; novel protein; tissue marker; molecular weight marker;	
XX	XW	chromosome marker; genetic disorder; gene; ds.	
XX	OS	Unidentified.	
XX	FN	WO2003054152-A2.	
XX	XX		

PD	03-JUL-2003.
XX	10-DEC-2002; 2002WO-US039555.
PF	10-DEC-2001; 2001US-0339739P.
XX	11-DEC-2001; 2001US-0339453P.
PR	14-MAR-2002; 2002US-0365091P.
PR	14-MAR-2002; 2002US-0365384P.
PR	12-APR-2002; 2002US-0372381P.
PR	12-APR-2002; 2002US-0372615P.
PR	22-APR-2002; 2002US-00128558.
PR	24-APR-2002; 2002US-0376045P.
XX	(HYSE-) HYSEQ INC.
PA	Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
XX	Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI	Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
PI	WPI; 2003-569235/53.
XX	P-PSDB; ADE08090.
DR	New polynucleotides, useful for expressing recombinant proteins for
DR	analysis, characterization or therapeutic use, or as markers for tissues
XX	in which the corresponding protein is preferentially expressed.
PT	Claim 1; SEQ ID NO 245; 1177pp; English.
PT	The invention comprises the amino acid and coding sequences of novel
XX	proteins. The DNA and protein sequences of the invention are useful as:
CC	markers for tissues in which the corresponding protein is preferentially
CC	expressed; as molecular weight markers on gels; as chromosome markers or
CC	tags; to identify chromosomes or to map related gene positions; and to
CC	compare with endogenous DNA sequences in patients to identify potential
CC	genetic disorders. The present DNA sequence represents a gene of the
CC	invention.
XX	Sequence 1534 BP; 462 A; 377 C; 355 G; 340 T; 0 U; 0 Other;
SQ	Alignment Scores:
	Pred. No.: 1.04e-53 Length: 1534
	Score: 975.00 Matches: 210
	Percent Similarity: 62.41% Conservative: 49
	Best Local Similarity: 50.60% Mismatches: 92
	Query Match: 45.73% Indels: 64
DB:	9 Gaps: 12
SEQ14-X-AT-28-64-76 (1-400) x ADE07179 (1-1534)	
QY	1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db	133 ATGGATGTGAAGAAGACGAGCCCTTACTGCTCCCTGACCAAGAGCAGACGAGAGGAA 192
QY	20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db	193 CGGCGCTACAAATTCCTCCGACAGACATGAGGAGTCCCGGGTACCCACACAGAGTCC 252
QY	39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db	253 TACAGTTCACGAGACATTTGAAAGCTTTTGATCATGATTCTCTCGCGGCTGCTTTACGGC 312
QY	58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db	313 AACAGAGTGAAGGATTGGTTTCACAGAGAGCAGACGAGTTCATAGACAGGACAGAAT 372
QY	78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db	373 TTACCCCTAAGGCAGTAGGAGTTTGTGAACCCAGCAACTCGAAGAGGACTGGCATTTTGT 432
QY	98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db	433 CGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTGACGGGTGATGCTGATACT 492

CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15b, which is
CC located on chromosome 4
XX
SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores: 7.52e-53 Length: 8645
Pred. NO.: 975.00 Matches: 210
Score: 62.41% Conservative: 49
Percent Similarity: 62.41% Mismatches: 92
Best Local Similarity: 50.60% Indels: 64
Query Match: 45.73% Gaps: 12
DB: 6

SEQ14-X-AT-28-64-76 (1-400) x ABQ82344 (1-8645)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAGAACGCGAGCCTTACTGCTCCCTGACCAAGAGCAGACGAGAGAGAA 210
QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGCGCTACACAAATTCTCCGACAGACAAATGAGGAGTGCCGGGTACCCACACAGAGATCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTCCAGCGAGACATTGAAGCTTTTGATCATGATTCTCGCGGTGCTTTACGGC 330
QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db 331 AACAGAGTGAAGGATTGTGTTTCACAGAGAAGCAGACGAGTTCTACTAGACAAGGACAGAA 390
QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTTAAGGCAGTTAGGAGTTTGTGAACCAAGCAACTCGAAGAGGACTGGCATTTTGT 450
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510
QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTATGATGTCCTCCAGAGCATGCCATGAGACTTTGGGCGAGGGGTCAA 570
QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCGGAGCTCTGCTGCTGTCAGTCCGTCCTCAACTCAGCCCTCAACNCTGACAGATACG 630
QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCACGAAACAAAGTCCGACAGTGAGAAAGACCAACCTGCAAGCAATCAAGGCCAGTCT 690
QY 175 ArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisAla 194
Db 691 ACCCTGCAGCCCTTGGCGCCT-----TCCCATAGCAGCACTCTGCACACATCAT--- 741
QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCTCAACAGAAACTCCCTGACCAATAGAAGAACCCAGAGTCCGGCC 801
QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGGCTGCTTGTGCTCCCGCGAGCTGCAAAACCA 837
QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCCGAGTCCGTCAGCTGCAGGACAGTGGTCCCTGGCAGTAATGTACCACCTGGAAGC 897
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269

Db 898 AGG----- 900
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTATTCAAACA 918
QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACACGCGCACTGTTTCAAGTACTGCAACCCAGGATACACAATGGCATCTGGC 978
QY 309 ThrValTyrSerProProArgProLeuProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTCACCACTACTCGGCCACTACCTAGAAAACACCCCTATCAAGAGTGTCTTT 1038
QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAAGAGTCTTCAAGTACTGTAGCTGGAATGCACTGCACTGTGTGCGTAGGG 1098
QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCCTGGCAATACTCCTGTCTTTATTTATAGCAATGCATCTCTTTGGCCTC 1158
QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAATGACACATTTGAGAATGGAAGTGAATTCGTGAT 1218
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1219 ACC-----ATGCCAACAAACACTGTGTGCTATACCT 1248
RESULT 11
ABQ82343
ID ABQ82343 standard; cDNA; 8675 BP.
XX
AC ABQ82343;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15a encoding cDNA SEQ ID NO:35.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurologic disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 151..8328
FT /*tag= a
FT /product= "NOV15a"
FT /transl_except= (pos:1249..1251,aa:Ser)
XX
PN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.

```

13-APR-2001; 2001US-0283889P.
18-APR-2001; 2001US-0284447P.
25-APR-2001; 2001US-0286683P.
29-MAY-2001; 2001US-0294080P.
16-AUG-2001; 2001US-0312915P.
17-AUG-2001; 2001US-0313325P.
17-SEP-2001; 2001US-0322699P.
26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53586.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 110-112; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4
XX
SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: Length: 8675
Score: 975.00 Matches: 210
Percent Similarity: 62.41% Conservative: 49
Best Local Similarity: 50.60% Mismatches: 92
Query Match: 45.73% Indels: 64
DB: 6 Gaps: 12
SEQ14-X-AT-28-64-76 (1-400) x ABQ82343 (1-8675)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
Db 151 ATGGATGTGAAGAAGACGACGGCCTTACTGCTCCCTGACCAAGACGACGAGAGAGAA 210
QY 20 ArgArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 211 CGCGGCTACACAATTCCTCCGACAGACAATGAGGAGTGCCGGGTACCCACACAGAGTCC 270
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57

```

```
RESULT 12
ACC72052
ID ACC72052 standard; DNA; 9695 BP.
XX
AC ACC72052;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205B gene #SEQ ID 81.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
PN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
WPI; 2003-381623/36.
DR P-PSDB; ABR58318.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.38e-50 Length: 9695
Score: 937.00 Matches: 206
Percent Similarity: 63.57% Conservative: 54
Best Local Similarity: 50.37% Mismatches: 93
Query Match: 43.95% Indels: 56
DB: 7 Gaps: 13

SEQ14-X-AT-28-64-76 (1-400) x ACC72052 (1-9695)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
Db 435 ATGGATGTAAAGGACCGCGCA---CACCGCTCTTTGACCAGAGGACGCTGTGGCAAAGAG 491
QY 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 492 TGTGCTACACAAGCTCTCTCTGGACAGTGAGGACTGCCCGGTGCCACACAGAAATCC 551
```

```
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 552 TACAGCTCCAGTGAGACTCTGAAGGCCTATGACCATGACAGCAGGATGCACTATGGAAC 611
QY 59 ArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPhe 78
Db 612 CGAGTCACAGACCTCATCCACCGGAGTCAGATGAGTTTCTTAGACAGGAACCACTTC 671
QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThr 98
Db 672 ACCCTTGCCGAACCTGGCATCTGTAGCCCTCC--CCACACCGAAGCGGTACTGTCTCC 728
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 729 GACATGGGATCCCTTCACAGGGCTACTCCCTTAGCACAGGGTCTGACGCCGACTCCGAC 788
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
Db 789 ACCGAGGAGGGATGCTCCAGAACACACGCCATCAGACTGTGGGCAGAGGGATAAATCC 848
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 849 AGGCGCAGTTCCGGCCTTCCAGTCTGTAACATCGGCCCTCCGAAACATCGGCTGACTGAC 908
QY 159 HisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArg-----Leu 176
Db 909 AACGAAACAAATCAGATGATGAGAACGGCCCTCCGAAACACACAGCCAGCTCGACTGTG 968
QY 177 ArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSer 196
Db 969 AGGCCCTCTCTCCACCC---CCTCACAACACACAGCTGTGCCATCACCAC---TCGTCC 1022
QY 197 IleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn-----ProSer 212
Db 1023 GCCAACTCCCTCAACAGGAACCTCACTGACCAATCGGGGAGTCTAGATCCAGCCCGGCC 1082
QY 213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro 232
Db 1083 CCAGGCCCAATGAC-----CTGGCCACCACACCA-----GAGTCC 1118
QY 233 AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu 252
Db 1119 GTTCAGCTTCAGGACAGCTGGTGTCTAAACAGCAACGTGCCACTGGAGACCCGG----- 1172
QY 253 GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly 272
Db 1172 ----- 1172
QY 273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys---ProGlyGly 291
Db 1173 -----CACTTCTCTTCAAGAGCTCCTCGGG 1199
QY 292 ThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyr 311
Db 1200 AGCACACCTTGTTTCAGCAGCTCTTCCCGGGATACCTTTGACCTCAGGAACGGTTTAC 1259
QY 312 SerProProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLys 331
Db 1260 AGCCCCCGCCCGCTGCTGCCAGGAATACTTTCTCCAGGAAGGCTTCAAGCTGAAG 1319
QY 332 LysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAla 351
Db 1320 AAGCCCTCCAAATACTGCAGCTGGAATGTGTGCCCTCTCCGCCATTGCCGGCCCTC 1379
QY 352 ThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHis 371
Db 1380 CTCTTGCTATTTGTGGCGTATTTTCATAGCAATGATCTGCTCGGACTCAATTGGCAA 1439
QY 372 LeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrpPro 391
Db 1440 CTCCAGCCTGCAGATGGGCACACCTTT-----AACATGGGATAAGGACCGGCTTACCA 1493
QY 392 ValProThrAspValSerLeuTyrPro 400
```


QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProProLeuSerHisAla 186
Db 924 CCTCCGAACACACAGCCAGTCGACTCTGAGGCCCTCTCCACCC---CCTCACAA 980
QY 187 HisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
Db 981 CACACGCTGCCATCACAC---TCGTCCGCCAACTCCCTCAACAGGAACACTGACC 1037
QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
Db 1038 AATCGCGGAGTCAGATCCAGCCCGGCCAGGCCCAATGAC-----CTGCCACC 1091
QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuAsn 242
Db 1092 ACACCA-----GAGTCCGTTTCAGCTTCAGGACAGCTGGGTCTAAAC 1133
QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
Db 1134 AGCAACGTGCCACTGGAGACCCGG-----1157
QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
Db 1157 -----1157
QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
Db 1158 ---CACTTCCTCTTCAAGACCTCCTCGGGAGCACACCCCTTGTTCAGCAGCTCTTCCCG 1214
QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProProLeuProArgSer 321
Db 1215 GGATACCTTTTGACCTCAGGAACGGTTTACAGCCCCCGCCCGCTGTGCCAGGAAT 1274
QY 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys 341
Db 1275 ACTTTCTCCAGGAAGGCTTTCAGTGAAGAACCCCTCCAATACTGCAGCTGGAATGT 1334
QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361
Db 1335 GCTGCCCTCTCGCCATTGCCCGCGCCCTCCTCTTGCTATTGTGCGGTATTTCATA 1394
QY 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln 373
Db 1395 -----GTGCCCTGTCGTTGAAA 1412
RESULT 14
ADB32028
ID ADB32028 standard; cDNA; 9729 BP.
XX
AC ADB32028;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human FCTR3f cDNA.
XX
KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
XX
OS Homo sapiens.
XX
PN US2003087816-A1.

XX 08-MAY-2003.
PD
XX
PF 05-MAR-2001; 2001US-00800198.
XX
PR 03-MAR-2000; 2000US-0186592P.
XX
PA (VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RAST/) RASTELLI L.
XX
PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2003-625633/59.
P-PSDB; ADB32029.
XX
PT New FCTR3 polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTR3-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
XX
PS Claim 9; Page 34-37; 155pp; English.
XX
CC The invention relates to FCTR3 polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTR3 polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, asthma, lung
CC conditions, immunological disorders, allergy and infection, deafness, liver
CC diseases, male and female reproductive disorders, diabetes, schistosoma
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, plasmodium falciparum infection,
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTR3 polypeptide of the invention.
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.42e-43 Length: 9729
Score: 828.00 Matches: 194
Percent Similarity: 53.98% Conservative: 50
Best Local Similarity: 42.92% Mismatches: 78
Query Match: 38.84% Indels: 130
DB: 9 Gaps: 14
SEQ14-X-AT-28-64-76 (1-400) x ADB32028 (1-9729)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
Db 210 ATGGATGTAAGGACCGCGCA---CACCGCTCTTGACCAGGACGCTGTGCAAGAG 266
QY 20 ArgArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 267 TGTGCTACACAAGCTCTCTCTGACAGTGGAGACTGCGGGTGCACACAGAATCC 326
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 327 TACAGCTCCAGTGAGACTCTGAAGGCCTATGACCATGACAGCAGGATGCATATGGAAC 386

QY 59 ArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPhe 78
Db 387 CGAGTACAGACCTCATCCACCGGAGTCAGATGATTTCTTAGACAGGAACCAACTTC 446
QY 79 ThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThr 98
Db 447 ACCCTGGCGAACTGGGCATCTGTAGCCCTCC--CCACACGAAGCGGCTACTGTCTCC 503
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 504 GACATGGGGATCCTTCACAGGGCTACTCCCTTAGCACAGGGTCTGACGCCGACTCCGAC 563
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSer 138
Db 564 ACCGAGGGAGGATGTCTCCAGAACACGCCCATCAGACTGTGGGCGAGGGATAAAATCC 623
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 624 AGGCGCAGTTCCGGCCTGTCCAGTCGTGAAGAACTCGGCCCTTACCCTGACTGACTCTGAC 683
QY 159 HisGluAsnThrGluThrAspHisProGly-----168
Db 684 AACGAAACAAATCAGATGATGAGAACGGTCTGCCATTCACCTACATCCTCGCCTAGT 743
QY 168 -----168
Db 744 CTCCTCCCATCTGCTCAGTCGCTAGCTCCCATATCTCCACCAGTTAGTGCCAGATG 803
QY 168 -----168
Db 804 CCATTGCTAGACAGCAACACCTCCCATCAATCATGGACACCAACCCTGATGAGGAATTC 863
QY 169 -----Gly 169
Db 864 TCCCCCAATTCTATACCTGCTCAGAGCATGCTCAGGGCCCCAGCAAGCCTCCAGCAGTGGC 923
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrProProProLeuSerHisAla 186
Db 924 CCTCCGAACCAACACAGCCAGTCGACTCTGAGGCCCCCTCTCCCACCC---CCTCACAAAC 980
QY 187 HisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
Db 981 CACACGCTGTCCCATCACCCAC---TCGTCCGCCCACTCCCTCAACAGGAACCTCACTGACC 1037
QY 207 ProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSerGly 222
Db 1038 AATCGGGGAGTCAGATCCAGCCCCCGGCCCAATGAC-----CTGGCCACC 1091
QY 223 GluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsn 242
Db 1092 ACACCA-----GAGTCCGTTTCAGCTTCAGGACAGCTGGGTGTAAAC 1133
QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
Db 1134 AGCAACGTGCCACTGGAGACCCCG-----1157
QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
Db 1157 -----1157
QY 283 GlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSerPro 301
Db 1158 ---CACTTCCTCTTCAAGACCTCCTCGGGAGCACACCCCTTGTTCAGCAGCTCTTCCCCG 1214
QY 302 GlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArgSer 321
Db 1215 GGATACCCCTTTGACCTCAGGAACGGTTTACACGCCCGCCCGCCCTGTGTCGCCAGGAAT 1274
QY 322 ThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCys 341
Db 1275 ACTTTCTCCAGGAAGGCTTCAAGCTGAAGAAGCCCTCCAAATACTGCAGCTGGAATGT 1334
QY 342 AlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal 361

Db 1335 GCTGCCCTCTCCGCCATTGCCGGCCCTCCTCTTGGCTATTGTGGCTATTTCATA 1394
QY 362 AlaMetHisLeuPheGlyLeuAsnTrpHisLeuGln 373
Db 1395 -----GTGCCCTGGTCGTTGAAA 1412
RESULT 15
AAS14085
ID AAS14085 standard; DNA; 9826 BP.
XX
AC AAS14085;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human FCTR3b DNA sequence.
XX
KW Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerbellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greonow type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW antiallergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3b; neurestin-like protein.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT 5'UTR 1..279
FT /*tag= b
FT CDS 280..8481
FT /*tag= a
FT /product= "Human FCTR3b"
FT 3'UTR 8482..9826
FT /*tag= c
XX WO200166747-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US007160.
XX
PR 03-MAR-2000; 2000US-0186592P.
PR 03-MAR-2000; 2000US-0186718P.
PR 06-MAR-2000; 2000US-0187293P.
PR 06-MAR-2000; 2000US-0187294P.
PR 17-MAR-2000; 2000US-0190400P.
PR 07-APR-2000; 2000US-0196018P.
PR 03-JAN-2001; 2001US-0259548P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2001-596837/67.
DR P-PSDB; AAU08680.
XX
PT Novel polypeptides designated as FCTRX polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
PS Claim 9; Page 33-35; 215pp; English.
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 120.435 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520
Perfect score: 372
Sequence: 1 TFWRSQVFIDHPVHLKFNV.....RLLTQEARSLGTPRQSRGT 71

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_1868 @runat_06082004_112215_29265 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372	100.0	2387	2 AAV19252	Aav19252 Human gam
2	372	100.0	3111	2 AAV19251	Aav19251 Human gam
3	372	100.0	3111	2 AAX87705	Aax87705 Gamma-her
4	372	100.0	3111	3 AAD00791	Aad00791 Human Her
5	372	100.0	3111	4 AAS18526	Aas18526 DNA encod
6	372	100.0	8354	6 ABS52100	Abs52100 Human TEN
7	372	100.0	8438	6 ABN85378	Abn85378 Human NOV
8	372	100.0	8645	6 ABS78652	Abs78652 Human CDN

9	254	68.3	8473	6	ABQ82345	Human NOV
10	254	68.3	8487	6	ABQ82346	Human NOV
11	254	68.3	8645	6	ABQ82344	Human NOV
12	234	62.9	8675	6	ABQ82343	Human NOV
13	220	59.1	12879	6	ABK92230	Prostate
14	220	59.1	13202	4	AAK51828	Human pol
15	195.5	52.6	1429	8	ACD40264	Human bre
16	194.5	52.3	1430	5	AAS14084	Human FCT
17	194.5	52.3	1430	9	ADB32021	Human FCT
18	194.5	52.3	1431	4	AAF27861	Human NOV
19	194.5	52.3	1727	7	ADA53142	Human cod
20	194.5	52.3	1743	5	AAS01213	Human cod
21	194.5	52.3	4245	7	AAL60066	DNA encod
22	194.5	52.3	9058	7	ACC72051	BCU0205A
23	194.5	52.3	9695	7	ACC72052	BCU0205B
24	194.5	52.3	9729	5	AAS14089	Human FCT
25	194.5	52.3	9729	9	ADB32028	Human FCT
26	194.5	52.3	9826	5	AAS14085	Human FCT
27	194.5	52.3	9826	9	ADB32023	Human FCT
28	166	44.6	1534	9	ADE07179	Novel cod
29	149	40.1	417	7	ACD05506	cDNA enco
30	149	40.1	643	5	AAS67419	DNA encod
31	149	40.1	1692	7	ACD05836	Novel hum
32	128	34.4	10242	4	ABL29075	Drosophil
33	128	34.4	17131	4	ABL29074	Drosophil
34	106	28.5	2569	4	ABL17563	Drosophil
35	70.5	19.0	1467	3	AAZ93796	Sequence
36	70.5	19.0	26776	2	AAX20254	Borrelia
37	69	18.5	3341	4	AAS27773	DNA encod
38	69	18.5	3341	9	ADB94576	Novel hum
39	69	18.5	3342	4	AAS27774	DNA encod
40	69	18.5	3342	9	ADB94577	Novel hum
41	68	18.3	2577	9	ADB63136	Human CDN
42	67.5	18.1	523	5	AAS05481	Mammalian
43	67.5	18.1	748	2	AAZ00430	Human sec
44	67.5	18.1	2105	4	ABL08203	Drosophil
45	67.5	18.1	4105	4	ABL08202	Drosophil

ALIGNMENTS

RESULT 1
AAV19252
ID AAV19252 standard; cDNA; 2387 BP.
XX
AC AAV19252;
XX
DT 17-AUG-1998 (first entry)
XX
DE Human gamma-hergulin cDNA clone 20.
XX
KW Gamma-hergulin; gamma-HRG; human; autocrine growth factor;
KW breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
KW cell proliferation; cell differentiation; cell survival;
KW neurological disorder; muscular disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 307..1890
FT /*tag= a
FT 1186..1263
FT /*tag= b
FT /*note= "insert DNA"
XX

WO9802541-A1.

22-JAN-1998.

08-JUL-1997; 97WO-US011841.

12-JUL-1996; 96US-0021640P.

PA (GETH) GENENTECH INC.
XX
PI Schaefer GM, Sliwkowski M;
XX
XX WPI; 1998-110589/10.
DR P-PSDB; AAW44818.
XX
XX DNA encoding gamma-hereregulin - used to activate ErbB receptor and to
PT enhance proliferation, differentiation or survival of a cell.
PT
XX
PS Disclosure; Page 56-57; 81pp; English.
XX
XX Clone 20 encodes an N-terminally truncated isoform (see AAW44818) of
CC human gamma-hereregulin (gamma-HRG) (see also AAW44817), a novel member of
CC the hereregulin superfamily. It was isolated from a MDA-MB-175 human breast
CC cancer cell library after screening with an EGF-like domain. Compared to
CC gamma-HRG cDNA (see AAV19251), clone 20 contains a 78 base insert between
CC codons 560 and 561, and lacks the 5' end. The truncated gamma-HRG was
CC insoluble when expressed as a thioredoxin fusion protein in Escherichia
CC coli cells. Gamma-HRG nucleic acid and polypeptide are claimed, together
CC with their uses e.g. for enhancing the proliferation, differentiation or
CC survival of glial or muscle cells
XX
SQ Sequence 2387 BP; 641 A; 640 C; 550 G; 556 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.43e-42 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAV19252 (1-2387)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 912

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 913 CTGGGAAAGGCAGCCCTGGTTGGCATTATGCGAGAAAGGCCTCCCTCTTCACATACA 972

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTAACCAGGAGCGCGGAGC 1032

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGGACCCGCGCCAGTCTCGGGGAACT 1065

RESULT 2
AAV19251
ID AAV19251 standard; cDNA; 3111 BP.
XX
AC AAV19251;
XX
DT 17-AUG-1998 (first entry)
XX
DE Human gamma-hereregulin cDNA.
KW Gamma-hereregulin; gamma-HRG; human; autocrine growth factor;
KW breast cancer; MDA-MB-175 cell; diagnosis; gene therapy; ErbB receptor;
KW cell proliferation; cell differentiation; cell survival;
KW neurological disorder; muscular disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 534..2645
FT /*tag= a
XX
PN WO9802541-A1.

XX 22-JAN-1998.
PD
XX
XX 08-JUL-1997; 97WO-US011841.
PF
XX
XX 12-JUL-1996; 96US-0021640P.
PR
XX (GETH) GENENTECH INC.
PA
XX Schaefer GM, Sliwkowski M;
PI
XX WPI; 1998-110589/10.
DR P-PSDB; AAW44817.
XX
PT DNA encoding gamma-hereregulin - used to activate ErbB receptor and to
PT enhance proliferation, differentiation or survival of a cell.
XX
PS Claim 21; Fig 1A-C; 81pp; English.
XX
CC This nucleic acid molecule codes for human gamma-hereregulin (gamma-HRG)
CC (see AAW44817), a novel member of the hereregulin superfamily, that has a
CC unique N-terminal domain not present in previously identified heregulins.
CC Gamma-HRG is secreted by human breast cancer MDA-MB-175 cells. It leads
CC to the formation of a constitutive active receptor complex and stimulates
CC the growth of these cells in an autocrine manner. The nucleic acid was
CC isolated by screening a MDA-MB-175 with a probe corresponding to the EGF-
CC like domain and part of the N-terminal sequences of HRG-beta-3. It can be
CC used for the recombinant production of gamma-HRG, or for in vivo or ex
CC vivo gene therapy. A claimed nucleic acid, which is complementary to the
CC nucleic acid sequence encoding the N-terminal domain of gamma-HRG, is
CC able to reduce production of gamma-HRG by MDA-MB-175 cells. The invention
CC provides claimed methods for activating an ErbB receptor and for
CC enhancing proliferation, differentiation or survival of a cell by
CC contacting the cell (preferably a glial or muscle cell) with gamma-HRG
XX
SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.03e-42 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAV19251 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThr 40
Db 1741 CTGGGAAAGGCAGCCCTGGTTGGCATTATGCGAGAAAGGCCTCCCTCTTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTAACCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCGCGCCAGTCTCGGGGAACT 1893

RESULT 3
AAV19251
ID AAV19251 standard; cDNA; 3111 BP.
XX
AC AAV19251;
XX
DT 26-OCT-1999 (first entry)
XX
DE Gamma-hereregulin cDNA.
XX

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 1740
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGGAAAGGCAGCCCTGTGGCATTTATGGCAGAAAGGCCTCCCTTCACATACA 1800
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlnSerArgGlyThr 60
Db 1801 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGTCTTAACCCAGGAGCGCGGAGC 1860
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGGAAC 1893

RESULT 5
AAS18526
ID AAS18526 standard; cDNA; 3111 BP.

AC AAS18526;

XX 26-FEB-2002 (first entry)

DE DNA encoding human heregulin, gamma-HRG.

XX Human; heregulin; antiasthmatic; antiinflammatory; vulnerary; antiulcer;
KW epithelial cell growth; HER2; HER3; HER4; HRG; lung cell; asthma;
KW lung surfactant protein A; respiratory distress; emphysema; pneumonitis;
KW chronic obstructive pulmonary disease; COPD; chronic bronchitis; ulcer;
KW neonatal pulmonary disease; neonatal respiratory distress syndrome;
KW meconium aspiration syndrome; congenital diaphragmatic hernia; lesion;
KW acute lung injury; cystic fibrosis; surgical wound; resection;
KW growth factor; smoke inhalation; gamma-HRG; ss.

XX Homo sapiens.

Key Location/Qualifiers
CDS 334..2640
FT /*tag= a
FT /product= "Human heregulin, gamma-HRG"

XX US2001023241-A1.

XX 20-SEP-2001.

XX 02-FEB-2001; 2001US-00773517.

XX 04-FEB-1998; 98US-0073866P.

XX 02-FEB-1999; 99US-00243198.

XX (SLIW/) SLIWKOWSKI M X.
XX (KERN/) KERN J A.

XX Sliwkowski MX, Kern JA;

XX WPI; 2001-595807/67.

XX P-PSDB; AAU09891.

PT Inducing epithelial cell growth and/or proliferation, useful in the
PT treatment of respiratory disease, comprises use of heregulin ligand as
PT growth factor.

PS Disclosure; Fig 7; 68pp; English.

XX The invention relates to inducing epithelial cell growth and/or
CC proliferation comprising contacting a normal epithelial cell, which
CC expresses HER2, HER3 and/or HER4 (preferably HER2/HER3, HER2/HER4,
CC HER3/HER4, HER3 or HER4) receptors with an isolated ligand such as
CC heregulin (HRG) which activates HER2, HER3 and/or HER4 receptors. The
CC method is used for inducing epithelial cell (preferably lung cell) growth
CC and/or proliferation, for increasing lung surfactant protein A and for

CC treating respiratory distress or emphysema, for treating chronic
CC obstructive pulmonary disease (COPD) e.g. chronic bronchitis, asthma,
CC neonatal pulmonary diseases including neonatal respiratory distress
CC syndrome, meconium aspiration syndrome, chronic lung disease of the
CC neonate, congenital diaphragmatic hernia and acute lung injuries
CC including smoke or chemical inhalation, pneumonitis due to aspiration,
CC radiation, near drowning, cystic fibrosis and other epithelial cell
CC trauma diseases including injuries associated with surgical wounds and
CC resections, ulcers, lesions and tissue tears as normal epithelial cell
CC growth factors. For treating infants/neonates with respiratory distress
CC as well as youth and adult with poor lung function due to lung injury or
CC damage. The HRG binds with varying and very high affinity to the HER2,
CC HER3 and/or HER4 receptors. The method stimulates growth and
CC proliferation of the epithelial cells, repairing and re-establishing the
CC cellular barriers of organs and allowing the affected tissue to develop
CC normal physiological functions more quickly. Hence the method improves
CC oxygenation and speeds of the development of a barrier to infection while
CC treating lung cells that are damaged by inhalation of smoke resulting in
CC emphysema. It also facilitates regeneration of epithelial cells. The
CC present sequence represents the coding sequence of human heregulin gamma-
CC HRG

SQ Sequence 3111 BP; 799 A; 873 C; 777 G; 662 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.03e-42 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AAS18526 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20

Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40

Db 1741 CTGGGAAAGGCAGCCCTGTGGCATTTATGGCAGAAAGGCCTCCCTTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlnSerArgGlyThrGlnGluAlaArgSer 60

Db 1801 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGTCTCTAACCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71

Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGGAAC 1893

RESULT 6

ABS52100

ID ABS52100 standard; DNA; 8354 BP.

XX AC ABS52100;

XX 05-NOV-2002 (first entry)

DE Human TEN-M4-like gene.

XX Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;

KW type 1a membrane sushi-containing domain; butyrophilin;
KW type 1a membrane-sushi domain containing; SNP; gene; ds;
KW single nucleotide polymorphism.

OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT variation replace(117,G)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT variation replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"

WO200257453-A2.

XX 25-JUL-2002.

XX 19-DEC-2001; 2001WO-US050331.

PR 19-DEC-2000; 2000US-0256704P.
PR 20-DEC-2000; 2000US-0257314P.
PR 02-MAY-2001; 2001US-0288153P.
PR 29-MAY-2001; 2001US-0294075P.
PR 24-JUL-2001; 2001US-0307506P.
PR 10-AUG-2001; 2001US-0311590P.
PR 10-AUG-2001; 2001US-0311613P.
PR 29-AUG-2001; 2001US-0315617P.
PR 14-SEP-2001; 2001US-0322358P.

XX (CURA-) CURAGEN CORP.

XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;

DR WPI; 2002-590744/63.

DR P-PSDB; ABG70388.

XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.

PS Claim 8; Page 50-52; 318pp; English.

XX The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders. The
CC present nucleic acid sequence represents a NOVX gene. This sequence
CC encodes a NOVX protein of the invention

SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.59e-42 Length: 8354
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABS52100 (1-8354)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1382 ACTTTCTGGAGATCTCAAGTGTTTCATAGACCATCTGTGCATCTGAAATTCAATGTGCT 1441

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1442 CTGGGAAAGGCAGCCCTGGTTGGCATTATGGCAGAAAGGCCTCCCTCCTTCACATACA 1501

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1502 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1561

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1562 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 1594

RESULT 7

ABN85378

ID ABN85378 standard; DNA; 8438 BP.

XX AC ABN85378;

DT 21-OCT-2002 (first entry)

DE Human NOV1, TEN-M4 like protein, coding sequence.

KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FH CDS 4..8395

FT /*tag= a

FT /trans_except= (pos: 1138..1147,aa:Met)

FT /product= "NOV1 protein"

XX WO200255704-A2.

XX 18-JUL-2002.

XX 09-JAN-2002; 2002WO-US000554.

XX 09-JAN-2001; 2001US-0260417P.

PR 10-JAN-2001; 2001US-0260831P.

PR 28-FEB-2001; 2001US-0272338P.

PR 09-MAR-2001; 2001US-0274876P.

PR 18-APR-2001; 2001US-0284704P.

XX (CURA-) CURAGEN CORP.

XX Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
PI Gunther E, Stone DJ;

XX WPI; 2002-590674/63.

DR P-PSDB; ABB98401.

XX NOVX polypeptides and encoding polynucleotides, useful for preventing or
PT treating NOVX-associated disorders e.g. cancer, inflammation, or
PT Alzheimer's disease, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 9; Page 8-9; 358pp; English.
XX
CC The present sequence is a coding sequence for a NOV protein. The NOV
CC proteins and coding sequences are useful for treating or preventing NOV-
CC associated disorders or in the manufacture of a medicament for treating
CC the disorders, such as cancer, heart disease, inflammation, autoimmune
CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
CC IgA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
CC and other wasting disorders associated with chronic diseases. NOV1 is a
CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX
SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.7e-42 Length: 8438
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABN85378 (1-8438)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1358 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1417

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1418 CTGGAAAGGCAGCCTGGTGGCATTTATGGCAGAAAGCCTCCCTCCATACATA 1477

QY 41 GlnPheAspPheValGlnLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1478 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGCTCCTAACCCAGGAGCGCGGAGC 1537

QY 61 LeuGlnGlyThrProArgGlnSerArgGlyThr 71
Db 1538 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1570

RESULT 8
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
AC ABS78652;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
KW Human; ss: gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200272830-A2.
XX

PD 19-SEP-2002.
XX
PF 08-FEB-2002; 2002WO-US003715.
XX
PR 09-FEB-2001; 2001US-0268111P.
PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
PR 09-MAR-2001; 2001US-0274552P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
DR WPI; 2002-723356/78.
DR P-PSDB; ABG97359.
XX
PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX
PS Claim 5; Page 175-178; 181pp; English.
XX
CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 7.95e-42 Length: 8645
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABS78652 (1-8645)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1452 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGCT 1511

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40

Db 1512 CTGGGAAGGCAGCCCTGGTTGGCATTATTATGGCAGAAAGGCCTCCCTTCACATACA 1571
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1572 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTTAACCCAGGAGCGCGGAGC 1631
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1632 CTAGAGGGGACCCCGCGCCAGTCTCGGGAACT 1664

RESULT 9
ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX AC ABQ82345;
XX DT 17-DEC-2002 (first entry)
XX DE Human NOV15c encoding cDNA SEQ ID NO:39.
XX KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 258..8144
FT /*tag= a
FT /product= "NOV15c"
XX PN WO200262999-A2.
XX PD 15-AUG-2002.

XX PF 31-DEC-2001; 2001WO-US049976.
XX PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.

XX PA (CURA-) CURAGEN CORP.
XX PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX WPI; 2002-732706/79.
DR P-PSDB; ABP53588.
XX PT New NOVX polypeptides and polynucleotides useful for treating NOVX-

PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.

XX Claim 8; Page 119-121; 444pp; English.

XX The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15c, which is
CC located on chromosome 4

XX SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 5.37e-25 Length: 8473
Score: 254.00 Matches: 48
Percent Similarity: 80.88% Conservative: 7
Best Local Similarity: 70.59% Mismatches: 13
Query Match: 68.28% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABQ82345 (1-8473)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1185 TTCTGGAGATCACAGCTCTTCATTGATCAGCCACAGTTTCTTAATTCAATATCTCTT 1244
QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
Db 1245 CAGAAAGGATGCATTGATTGGAGTATATGGCCGGAAGGCTTACCGCTTCCCATCTCAG 1304
QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61
Db 1305 TATGACTTCGTGGAGCTCCTGGATGGCAGCAGGCTGATTGCCAGAGACGAGCGGCTG 1364
QY 62 GluGlyThrProArgGlnSerArg 69
Db 1365 CTTGACAGCGGAGAGAGCCCGGCGG 1388

RESULT 10

ABQ82346
ID ABQ82346 standard; cDNA; 8487 BP.

XX AC ABQ82346;

XX DT 17-DEC-2002 (first entry)

XX DE Human NOV15d encoding cDNA SEQ ID NO:41.

XX KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;

KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

OS Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT CDS 299..8140
FT /*tag= a
FT /product= "NOV15d"

XX
PN WO200262999-A2.
XX
XX
PD 15-AUG-2002.
XX
XX 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.
DR P-PSDB; ABP53589.

XX
XX
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.

XX Claim 8; Page 123-125; 444pp; English.

XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be

CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15d, which is
CC located on chromosome 4

XX Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;
SQ

Alignment Scores:
Pred. No.: 5.38e-25 Length: 8487
Score: 254.00 Matches: 48
Percent Similarity: 80.88% Conservative: 7
Best Local Similarity: 70.59% Mismatches: 13
Query Match: 68.28% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABQ82346 (1-8487)

QY 2 PheTTPArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1226 TTCTGGAGATCACAGCTCTTCATHGATCAGCCACAGTTTCTTAAATCAATATCTCTT 1285

QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProSerHisThrGln 41
Db 1286 CAGAAAGGATGCATTGATTGGAGTATATGCCGGAAGGCTTACCGCCTTCCCATACTCAG 1345

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61
Db 1346 TATGACTTCGTGGAGCTCCTGGATGGCAGCAGGCTGATTGCCAGAGCAGCGGAGCCTG 1405

QY 62 GluGlyThrProArgGlnSerArg 69

Db 1406 CTTGAGACGGAGAGACCGGGCGG 1429

RESULT 11

ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.

XX ABQ82344;

XX 17-DEC-2002 (first entry)

DE Human NOV15b encoding cDNA SEQ ID NO:37.

XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 151..8316
FT /*tag= a
FT /product= "NOV15b"

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.

PR 02-JAN-2001; 2001US-0259415P.

PR 04-JAN-2001; 2001US-0259785P.

PR 20-FEB-2001; 2001US-0269814P.

PR 09-MAR-2001; 2001US-0279863P.

PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 110-112; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytosstatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiviral, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4
XX
SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 3.97e-22 Length: 8675
Score: 234.00 Matches: 47
Percent Similarity: 75.00% Conservative: 7
Best Local Similarity: 65.28% Mismatches: 14
Query Match: 62.90% Indels: 4
DB: 6 Gaps: 1

US-10-029-020-14_COPY_450_520 (1-71) x ABQ82343 (1-8675)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1357 TTCTGGAGATCACAGCTCTTCATTGATCAGCCACAGTTTCTTAAATTCATATCTCTTT 1416
QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr--- 40
Db 1417 CAGAAGGATGCATTGATTGGAGTATATGGCCGGAAGAAGTTACCGCTTCCCATACTCAG 1476
QY 41 -----GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGlu 57
Db 1477 TCCTCCCCCAGTATGACTTCGTGGAGCTCCTGGATGGCAGCAGGCTGATTGCCAGAGAG 1536
QY 58 AlaArgSerLeuGluGlyThrProArgGlnSerArg 69
Db 1537 CAGCGGAGCCTGCTTGACGCGGAGAGAGCCGGCGG 1572

RESULT 13
ABK92230
ID ABK92230 standard; DNA; 12879 BP.
XX
AC ABK92230;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated DNA sequence #116.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene; ds.
XX
OS Mammalia.
XX
PN WO200230268-A2.

XX 18-APR-2002.
PD
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
XX WPI; 2002-471335/50.
DR P-PSDB; ABG61913.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
PS Claim 22; Page 394-397; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences
XX
SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.69e-20 Length: 12879
Score: 220.00 Matches: 36
Percent Similarity: 79.41% Conservative: 18
Best Local Similarity: 52.94% Mismatches: 14
Query Match: 59.14% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x ABK92230 (1-12879)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 1313 TTCTGGCGTTTCCAGATTACTATCCACCATCCATATATCTGAAGTTCAATATTTCTTTA 1372
QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
Db 1373 GCCAAGGACTCTCTGCTGGGAATTTATGGCAGAGAAGAACATTCCACCTACACATACTCAG 1432
QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61
Db 1433 TTGATTTTGTAAACTAATGATGGCAACAGCTGGTCAAGCAGGACTCCAAGGGCTCT 1492
QY 62 GluGlyThrProArgGlnSerArg 69
::: |||

Db 1493 GATGATACACAGCACTCCCTCGG 1516

RESULT 14

AAK51828

ID AAK51828 standard; cDNA; 13202 BP.

XX

AC AAK51828;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 373.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorder; arthritis; inflammation; ss.

XX

OS Homo sapiens.

XX

PN WO200157190-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US004098.

XX

PR 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.

PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.

PR 30-NOV-2000; 2000US-00728422.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

DR WPI; 2001-476283/51.

DR P-PSDB; AAM78695.

XX

PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

PT in diagnosis and gene therapy.

XX

PS Claim 1; Page 1414-1426; 6221pp; English.

XX

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

CC sequence listing were missing at the time of publication

XX

SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.91e-20	Length:	13202
Score:	220.00	Matches:	36
Percent Similarity:	79.41%	Conservative:	18
Best Local Similarity:	52.94%	Mismatches:	14
Query Match:	59.14%	Indels:	0
DB:	4	Gaps:	0

US-10-029-020-14_COPY_450_520 (1-71) x AAK51828 (1-13202)

QY 2 PheTIpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21

Db 1639 TTCTGGCGTTTCCAGATTACTATCCACCATCCAAATATATCTGAAGTTCAATTCTTTA 1698

QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41

Db 1699 GCCAAGGACTCTCTGCTGGGAATTTATGGCAGAGAAACATTCCACCTACACATACTCAG 1758

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61

Db 1759 TTTGATTTTGTAAAACTAATGGATGGCAACACAGCTGGTCAAGCAGGACTCCAAGGGCTCT 1818

QY 62 GluGlyThrProArgGlnSerArg 69

Db 1819 GATGATACACAGCACTCCCTCGG 1842

RESULT 15

ACD40264

ID ACD40264 standard; DNA; 1429 BP.

XX

AC ACD40264;

XX

DT 03-SEP-2003 (first entry)

XX

DE Human breast tumour associated protein 47-like polypeptide NOV11 DNA.

XX

KW Tissue typing; cancer; breast cancer; colon cancer; lung cancer; sarcoma;

KW pancreatic cancer; uterine cancer; organ transplantation disorder; ds;

KW cardiovascular disease; melanoma; atherosclerosis; diabetes mellitus;

KW ischaemic heart disease; haemorrhage; peripheral vascular disease; gene;

KW thrombosis; hypertension; systemic lupus erythematosus; haematopoiesis;

KW tissue regeneration; wound healing; hyperproliferative disorder;

KW psoriasis; neural disorder; Parkinson's disease; Alzheimer's disease;

KW Huntington's disease; multiple sclerosis; amyotrophic lateral sclerosis;

KW ALS; peripheral neuropathy; nervous system tumour; neurotoxin; tremor;

KW neuropathy; acute brain injury; peripheral nerve trauma; human; NOVX;

KW gene therapy; epilepsy; breast tumour associated protein 47.

XX

OS Homo sapiens.

XX

PN US2003027158-A1.

XX

PD 06-FEB-2003.

XX

PF 15-OCT-2001; 2001US-00977418.

XX

PR 03-JUN-1999; 99US-0137322P.

PR 16-MAR-2000; 2000US-0189810P.

PR 22-MAR-2000; 2000US-0191158P.

PR 30-MAR-2000; 2000US-0193086P.

PR 03-MAY-2000; 2000US-0201388P.

PR 31-MAY-2000; 2000US-00584411.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Shimkets RA, Fernandes E, Herrman J, Vernet C;

XX

DR WPI; 2003-492028/46.

DR P-PSDB; ABO23243.

XX

PT New nucleic acid sequence encoding a human breast tumor-associated

PT protein 47-like polypeptide, useful for treating cardiovascular

PT disorders, neural disorders, diabetes mellitus and cancers.

XX

PS Disclosure; Page 29-30; 100pp; English.

XX

CC The invention relates to a new isolated NOV4 nucleic acid. The nucleic

CC acid is useful for identifying a compound that binds the nucleic acid.

CC The nucleic acid is useful in gene therapy, in screening assays, in

CC detection assays e.g. chromosomal mapping, cell and tissue typing and

CC forensic biology, predictive medicine e.g. diagnostic assays, prognostic

CC assays, monitoring clinical trials and pharmacogenomics and methods of

CC treatment including therapeutic and prophylactic. The nucleic acid is
CC also useful for expressing NOVX protein. The nucleic acid is also useful
CC to provide polynucleotide reagents e.g. labelled probes that are useful
CC in an in situ hybridisation technique, for identifying a specific tissue
CC (for example brain tissue) and for use in forensic science. The nucleic
CC acid is also useful for mapping genes on a chromosome and thus locating
CC gene regions associated with genetic disease, identifying an individual
CC from a minute biological sample and to aid in forensic identification of
CC biological sample. The nucleic acid is also useful for treating cancer,
CC especially cancers of the breast, colon, lung, pancreas or uterus, or a
CC melanoma or sarcoma. The nucleic acid is also useful for treating
CC disorders related to organ transplantation, cardiovascular diseases,
CC atherosclerosis, ischaemic heart disease, haemorrhage, diabetes mellitus,
CC peripheral vascular disease, thrombosis, hypertension and systemic lupus
CC erythematosus. NOVX protein encoded by the nucleic acid is useful for
CC regulating haematopoiesis, for regeneration of bone, cartilage, tendon
CC ligament and/or nerve tissue growth or regeneration and for wound
CC healing. The nucleic acid is also useful for treating infections,
CC hyperproliferative disorders e.g. psoriasis, and neural disorders
CC including Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral
CC neuropathy, tumours of the nervous system, exposure to neurotoxins, acute
CC brain injury, peripheral nerve trauma or injury and other neuropathies,
CC epilepsy, and/or tremors. The present sequence represents DNA encoding a
CC human breast tumour associated protein 47-like polypeptide

XX
SQ Sequence 1429 BP; 349 A; 426 C; 340 G; 312 T; 0 U; 2 Other;

Alignment Scores:

Pred. No.:	1.12e-17	Length:	1429
Score:	195.50	Matches:	38
Percent Similarity:	70.59%	Conservative:	10
Best Local Similarity:	55.88%	Mismatches:	15
Query Match:	52.55%	Indels:	5
DB:	8	Gaps:	1

US-10-029-020-14_COPY_450_520 (1-71) x ACD40264 (1-1429)

QY	2	PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu	21
Db	884	TTTTGGAGGTCACAAATTCACATCAGTCAGCCCCAGTCTTAAAGTTCAACATCTCCCTC	943
QY	22	GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln	41
Db	944	GGGAAGGACGCTCTCTTTGGTGTTTACATAAGAACAGGACTTCCACCATCTCATGCCCAG	1003
QY	42	PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu	61
Db	1004	TATGACTTCATGGAACCTCTGGACGGGAAG-----GAGNAGTGGAGTGTG	1048
QY	62	GluGlyThrProArgGlnSerArg	69
Db	1049	GTTGAGTCTCCCGAAGGACGCCGG	1072

Search completed: August 14, 2004, 02:28:55
Job time : 135.435 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: August 13, 2004, 23:35:57 ; Search time 1786.37 Seconds
(without alignments)
2450.585 Million cell updates/sec
Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WMGAACDQACHPRCAEHGT.....DTSMETACGDSKNDGDGLV 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29:	em_vi:*
30:	em_htg_hum:*
31:	em_htg_inv:*
32:	em_htg_other:*
33:	em_htg_mus:*
34:	em_htg_pln:*
35:	em_htg_rod:*
36:	em_htg_mam:*
37:	em_htg_vrt:*
38:	em_sy:*
39:	em_htgo_hum:*
40:	em_htgo_mus:*
41:	em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	617	100.0	8354	6	AX556500	AX556500 Sequence
2	617	100.0	8645	6	AX600210	AX600210 Sequence
3	605	98.1	8585	10	AB025413	AB025413 Mus muscu
4	579	93.8	8438	6	AX675551	AX675551 Sequence
5	541	87.7	8964	10	AB025412	AB025412 Mus muscu
6	540.5	87.6	9722	10	AF059485	AF059485 Mus muscu
7	515	83.5	8473	6	AX662357	AX662357 Sequence
8	515	83.5	8645	6	AX662355	AX662355 Sequence
9	515	83.5	8675	6	AX662353	AX662353 Sequence
10	503.5	81.6	9264	5	AB026980	AB026980 Danio rer
11	483.5	78.4	3394	9	AK125869	AK125869 Homo sapi
12	483.5	78.4	7816	10	AF195418	AF195418 Mus muscu
13	478.5	77.6	8487	6	AX662359	AX662359 Sequence
14	467.5	75.8	8373	10	AB025410	AB025410 Mus muscu
15	464.5	75.3	8297	9	AF100772	AF100772 Homo sapi
16	454.5	73.7	8816	5	AB026979	AB026979 Danio rer
17	454	73.6	816	10	AF086610	AF086610 Rattus no
18	454	73.6	1011	10	AF086609	AF086609 Rattus no
19	454	73.6	7781	9	AB032953	AB032953 Homo sapi
20	454	73.6	9826	6	AX250008	AX250008 Sequence
21	452	73.3	8409	5	GGA279031	AJ279031 Gallus ga
22	452	73.3	8409	6	AX250068	AX250068 Sequence
23	448.5	72.7	7713	9	HSM808325	EX648178 Homo sapi
24	448.5	72.7	8689	6	AX250067	AX250067 Sequence
25	448.5	72.7	8689	10	AF086607	AF086607 Rattus no
26	448.5	72.7	8797	6	AX250063	AX250063 Sequence
27	448.5	72.7	8797	6	AX250066	AX250066 Sequence
28	448.5	72.7	8797	10	AB025411	AB025411 Mus muscu
29	448.5	72.7	9729	6	AX250013	AX250013 Sequence
30	447.5	72.5	8575	6	AX921803	AX921803 Sequence
31	435.5	70.6	8118	5	GGA238613	AJ238613 Gallus ga
32	330.5	53.6	798	10	AF086608	AF086608 Rattus no
C 33	301	48.8	112872	2	AP000716	AP000716 Homo sapi
C 34	301	48.8	165812	2	AP001141	AP001141 Homo sapi
C 35	301	48.8	168308	2	AC084775	AC084775 Homo sapi
C 36	301	48.8	170156	2	AP002412	AP002412 Homo sapi
C 37	301	48.8	175464	2	AP002887	AP002887 Homo sapi
C 38	301	48.8	186084	9	AP002768	AP002768 Homo sapi
C 39	300	48.6	245722	2	AC120288	AC120288 Rattus no
C 40	300	48.6	257693	2	AC103323	AC103323 Rattus no
41	296	48.0	10471	3	AF008228	AF008228 Drosophil
42	296	48.0	11451	3	DMTENM	X73154 D.melanogas
C 43	290	47.0	226256	10	AC073599	AC073599 Mus muscu
44	279.5	45.3	10216	3	DMTENA	X68794 Drosophila
45	278	45.1	14045	3	DMODZ4	AF008227 Drosophil

ALIGNMENTS

Mon Aug 16 09:01:16 2004

us-10-029-020-14_copy_750_850.p2n.rge

AX556500 LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
Zernhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,
Smithson,G., Li,L. and Ji,W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source
1. .8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 6.06e-43 Length: 8354
Pred. No.: 617.00 Matches: 101
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-029-020-14_COPY_750_850 (1-101) x AX556500 (1-8354)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2282 TGGATGGGGCAGCCTGCGACACGAGCGGCGCTGCCACCGCGCTGTGCCGAGCATGGGACC 2341
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2342 TGCCGCGACGGCAAGTCGAGTGCAGCCCTGGCTGGTAATGGCAACACTGCACCATCGCT 2401
QY 41 HisTyrLeuAspArgValLysGluCysProGlyCysProGlyLeuCysAsnGlyArg 60
Db 2402 CACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGCTGGTAATGGCAACACTGCACCATCGCT 2461
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2462 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGGC 2521
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2522 TGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGACACATGATGGAGATGGCCTG 2581
QY 101 Val 101
Db 2582 GTG 2584
RESULT 2
AX600210 8645 bp DNA linear PAT 14-FEB-2003
LOCUS
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,

Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Walia,N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7488573CB1"

ORIGIN

Alignment Scores: 6.26e-43 Length: 8645
Pred. No.: 617.00 Matches: 101
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-029-020-14_COPY_750_850 (1-101) x AX600210 (1-8645)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2352 TGGATGGGGCAGCCTGCGACACGCGGCGCTGCCACCGCGCTGTGCCGAGCATGGGACC 2411
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2412 TGCCGCGACGGCAAGTCGAGTGCAGCCCTGGCTGGTAATGGCGAACACTGCACCATCGCT 2471
QY 41 HisTyrLeuAspArgValLysGluCysProGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2472 CACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGCTGGTAATGGCAACACTGCACCATCGCT 2531
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2532 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGGC 2591
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2592 TGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGACACATGATGGAGATGGCCTG 2651
QY 101 Val 101
Db 2652 GTG 2654

RESULT 3

AB025413
LOCUS
DEFINITION Mus musculus mRNA for Ten-m4, complete cds.
ACCESSION AB025413
VERSION AB025413.1 GI:4760781
KEYWORDS Ten-m4.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (sites)
AUTHORS Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
TITLE Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
JOURNAL J. Cell Biol. (1999) In press
REFERENCE 2 (bases 1 to 8585)
AUTHORS Oohashi,T.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8522, Japan
(E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)
Location/Qualifiers

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

US-10-029-020-14_COPY_750_850 (1-101) x AB025413 (1-8585)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
 Db 2441 TGGATGGGGCCGCATGCGACCAACGGGCTGCCACCACCGCTGTGCAGAACACGGGACC 2500

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
 Db 2501 TGCCGGGACGGCAAGTGCGAATGCAGCCCCGGCTGGAATGGAGAGCACTGCACCATCGCT 2560

QY 41 HisTyrLeuAspArgValValLysGluCysProGlyLeuCysAsnGlyAsnGlyArg 60
 Db 2561 CACTATCTGGATAGGTAGTTAAAGAGGGCTGTCTGGCTTGTGCAATGGAAATGGCAGA 2620

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
 Db 2621 TGTACCTGGACCTGAATGGGTGGCACTGTCTGCCAGCTGGCTGGCGGAGGACTGGC 2680

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
 Db 2681 TGCAGACATCCATGGAAACGGGCTGTGGAGACGGCAAGGACAACGACGAGATGGCTTG 2740

QY 101 Val 101
 Db 2741 GTG 2743

RESULT 4
 AX675551

LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 1 from Patent WO02055704.
 ACCESSION AX675551
 VERSION AX675551.1 GI:293333552
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Padigaru, M., Li, L., Zerhusen, B.D., Casman, S.J., Shenoy, S.,
 Spytek, K.A., Zhong, M., Gangolli, E.A., Burgess, C.E., Patturajan, M.,
 Vernet, C.A., Taylor, S., Tchernev, V.T., Miller, C.E., Guo, X.,
 Boldog, F.L., Grosse, W.M., Alsobrook, J.P., Gerlach, V.,
 Edingermark, S., Rotherberg, M.E., Ellerman, K., Macdougall, J.,
 Malyankar, U., Millet, I., Peyman, J., Smithson, G., Gunther, E. and
 Stone, D.J.
 TITLE Proteins, polynucleotides encoding them and methods of using the
 same
 JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
 Curagen Corporation (US)
 FEATURES
 source 1. .8438
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 1.06e-39 Length: 8438
 Score: 579.00 Matches: 98
 Percent Similarity: 94.29% Conservative: 1
 Best Local Similarity: 93.33% Mismatches: 2
 Query Match: 93.84% Indels: 4
 DB: 6 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x AX675551 (1-8438)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
 Db 2249 TGGACGGGCCAGCTGTAATCAGAGAGCCTGCCACCCCCCGCTGTGCCGAGCACGGGACC 2308

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
 Db 2561 CACTATCTGGATAGGTAGTTAAAGAGGGCTGTCTGGCTTGTGCAATGGAAATGGCAGA 2620

ORIGIN

Alignment Scores:		
Pred. No.:	6.55e-42	8585
Score:	605.00	98
Percent Similarity:	97.03%	Conservative: 0
Best Local Similarity:	97.03%	Mismatches: 3
Query Match:	98.06%	Indels: 0
DB:	10	Gaps: 0

Db 2309 TGCCGCGACGGCAAGTCGAGCTGCAGCCCTGGCTGGATGGCGAACAACCTGCACCATCTCC 2368

Qy 40 ---AlaHisTyrLeuAspArgValVallys-----GluGlyCysProGlyLeuCysAsn 56

Db 2369 CTAGCTCACTATCTGGATAGGTAGTTAAACTTTTCAGAGGGTTGCCCTGGTTGTGCAAT 2428

Qy 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76

Db 2429 GGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGG 2488

Qy 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAsp 96

Db 2489 AGAGGAGCTGGCTGTGACACTTCATGGAGACTGCCTGCGGTGACAGCAAGACAATGAT 2548

Qy 97 GlyAspGlyLeuVal 101

Db 2549 GGAGATGGCCTGGTG 2563

RESULT 5

AB025412

LOCUS

DEFINITION Mus musculus mRNA for Ten-m3, complete cds.

ACCESSION AB025412

VERSION AB025412.1 GI:4760779

KEYWORDS Ten-m3.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (sites) Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T., Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.

AUTHORS Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues

TITLE J. Cell Biol. (1999) In press

JOURNAL 2. (bases 1 to 8964)

REFERENCE Oohashi,T.

AUTHORS Direct Submission

TITLE Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University

JOURNAL Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho, Okayama, Okayama 700-8252, Japan

(E-mail: oohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128, Fax:+81-86-222-7768)

FEATURES

source

1. .8964

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Balb/c"

/db_xref="taxon:10090"

/tissue_type="brain"

/dev_stage="adult"

1. .8964

/gene="ten-m3"

50. .8197

/gene="ten-m3"

/codon_start=1

/product="Ten-m3"

/protein_id="BAA77398.1"

/db_xref="GI:4760780"

/translation="MDVKERRPYCSLTKSRRREKERRYNTSSADNEECRVPTQKSYSSS

ETLKAFDHDYSRLLYGNRVKDLVHREADEYTRQGNFTLRQLGVCESATRRGVAFCAE

MGLPHRGYSISAGSDADTENEAVMSPEHAMRLWGRVKPSGSRSSLSRNSALTLTDT

EHENRSDSESQSPNNPGQPTLQPLPPSHKHQHPAQHHPSITSLNRNSLTNRNNOQSPAP

PAALPAELQTTPESVQLQDSWVLGSNVPLESRHFLFKTGTGTTPLFSTATPGYTMASG

SVYSPPTPLPRLNTLSRSFAFKFKSKYCSWRCTALCAVGSVSLAILLSYFIAMHLF

GLNWHLQOTENDTFFENGKVNSTVPTNTVSLPSGDNGLKGGFTHENNTIDSGELDIGR

RAIQEVFPGIFWRSOLFIDQPQFLKFNISLQDALIGVYGRKGLPPSHTQYDFVELLD

GSRLIAREQNLVSESRAGROARSVSLHEAGFIQYLDGSIWHLAFYNDGKNPEQVSFN

TIVIESVVECPRNCHNGECVSGTCHCFPGFLGPDCSRACPVLCSGNGQYSKGRCLC

FSGWKGTEDCVPTTQCIDPQCGGRI CIMGSCACNSGYKGENCEADCLDPGCSNHGV

CIHGECHNPGWGGSNCEILKTMCAQCSGHGTYLQESGCTCDPNWTGPDSCNEICS

VDCGSHGVMGGSRCRCEGWTGPACNORACHPRCAEHGTCCKGCECSQGWNGEHCTI

AHYLDKIYKEGCPGLCNSNGRCTLDQNGWHCVCPQGWRGAGCDVAMETLCTDSKDNEG

DGLIDCMPDCCLOSSCQNPYCRGLPDQDIIISQLQTPSQQAAKSFYDRISFLIGS

DSTHVLPGESPENKSLASVIRGOVLTAADGTPLIGVNVSVFLHYSEYGYTITRODMFDL

VANGGASPLTLFERSPFLTQYHTVMIPWNFYVMDTLVMKKEENDIPSCDLSGFVRPS

PIIVSSPLSTFRSSPEDSPIIPETQVLHEETIPTGDLKLSYLSSRAAGYKSVLKIT

MTQAVIPFNLMMKVLMAVAVGRLFQKWFPAFNLAYTFIWDKTDAYNQKVGLSEAVV

SVGYEYESCLDLTLWEKRTAVLQGYELDASNMGGWTLDKHHVLDVQNGILYKNGENQ

FISQOPPVSIMNGRRRSISPCSCNGQADGNKLLAPVALACGIDGSLYGVGFNYVR

RIFSPGNVTSVLELRNKDFRHSSNPAHRYYLADVTGDLVYSDTNTNRIIRPKSLTG

AKDLTKNAEVVAGTGEQCLPFDEARCGDGGKAVTEATLMSPKGMAIDKNGLIYFVDTGM

IRKVDQNGIIISTLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDN

NVVLQITENRQVRIAAGRPMHCQVPGVEYVPGKHAQTTLLESATAIAVSYSGVLYITE

TDEKKINRIROVTTDGEISLVAGIPSECDCKNDANDCYQSGDGYAKADKLNAPSSLA

ASPDGTLYIADLGNIRIRAVSKNKLNSMNFVEVASPTDQELYIFDINGHQYTVSL

VTGDLYNFSYSDNDVTAVTDSNGNTLRIIRDNRMPVRVVSQNVQIWLITGTNGC

LKSMTAQGLELVLFTHGNSGLLATKSDETGTTFFDYDSEGRLTNTVFTPTGVVTNLH

GDMDKAITVDIESSSREEDVSIITSNLSSIDSFTVMVQDLRNSYQIGYDGLRIFYAS

GLDSHYQTEPHVLAGTANPTVAKRNTLPGENGONLVEWFRKEQAQGXNVFGRKLR

VNGRNLSSVDFDRTTKTEKIYDDHRKFLRLRIAYDTSGHPTLWLPSSKLMVNVYTSST

GQIASIQRTTSEKVDYSQGRIVSRVFADGKTWSYTYLEKSMVLLHLSQRQYIPEYD

MWDRLSAITMPSVARHTMQTIRSIGYRNIYNPPESNASIITDYNEEGLLLOTAFLGT

SRVLFKRYRRQTRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFCITIRYRQIGPLI

DRQIFRSEEDGMVNARFDYSYDNSFRVTSMQGVNETPLPIDLYQFDDISGKVEQFGK

FGVIYYDINQIIISTAWMTYTKHFDAGHRIKEIQEIEFRSLMWITIQYDNMGRVTKRE

IKIGPFANTTKYAYEYDVGQLQTVLYLNEKIMWRVNYDNLGNLHLLNPSSSARLTPLR

YDLRDRITRLGDVOYRLDEDDGFLRQRGTEIFEYSKGLLTRVYSKSGWTVIYRDGL

GRRVSSKTSLGQHLFFYADLTYPTRITHVYHNSSEITSLYDLDLQHLFAMEIISGD

EFVIASDNTGTPLAVFSSNGLMKQTYTAYGEIYFSDNSVDQFVIGFHHGLYDPLTK

LIHFGERDYDILAGRWTPDIEIWKRIKDPAPFNLYMFRNNNPASKIHDVKDITDV

NSWLVTFGHLHNAIPGPVPKFDLTSPSYELVKSQOWEDVPPIFGVQQQVARQAKAF

LSLGKMAEVQVRRKAGAEQSWLWFATVKSLSIGKVMLAVSQGRVQTNVLNIANEDCI

KVAAVLNNAFYLENLHFTIEGKDTHYFIKTTTPESDLGTLRLTSGRKALENGINTVS

QSTTVVNGRTRRFADVEMQFGALALHVRYGMTLDEKARILEQARQALARAWAREQQ

RVRDGEEGARLWTEGEKREQLLSAGKVQGYDGYVYVLSVEQYPELADSANNIQFLRQSEI

GKR"

ORIGIN

Alignment Scores:

Pred. No.: 1.95e-36 Length: 8964

Score: 541.00 Matches: 85

Percent Similarity: 91.09% Conservative: 7

Best Local Similarity: 84.16% Mismatches: 9

Query Match: 87.68% Indels: 0

DB: 10 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x AB025412 (1-8964)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20

Db 2153 TGGACCGGCGCGCGTGTATCAGAGAGCTTGCCACCCTCGCTGCTGAGCACGGGACG 2212

Qy 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40

Db 2213 TGCAAGGACGGCAAGTCCGAGTGCAGCCCAAGGATGGAACGGAGAGCACTGCACAATTGCT 2272

Qy 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60

Db 2273 CACTATTTGGATAAGATAGTTAAAGAGGGTTGCCCGCGCTTGTGCAACAGCAATGGGAGA 2332

Qy 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80

Db 2333 TGCACACTGGACCAAAACGGCTGGCACTGCGCTTGCCAGCCAGGGTGGAGAGGAGCAGGC 2392

Qy 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100

Db 2393 TGTGACGTAGCCATGGAGACCCCTCTGTACAGACAGCAAAAGACAACGAGAGAGACGACTC 2452

Qy 101 Val 101

Db 2453 ATT 2455

RESULT 6

AF059485

LOCUS

9722 bp mRNA linear ROD 15-AUG-1998

DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 9722)
Wang,X.Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P., Zinszner,H. and Ron,D.
Identification of novel stress-induced genes downstream of chop EMBO J. 17 (13), 3619-3630 (1998)
98315054
9649432
PUBMED
REFERENCE 2 (bases 1 to 9722)
Wang,X.-Z. and Ron,D.
Direct Submission
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
JOURNAL
FEATURES
source
1. .9722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1. .9722
/gene="Doc4"
/note="downstream of CHOP 4"
583. .9060
/gene="Doc4"
/note="similar to Drosophila melanogaster tenm/odz and human gamma-herregulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translation="MDVKERKPYRSLTRRRDAERYTSSADSEEGKPKQSYSSSET LKAYDQDARLAYGSRVKDMVPQEAEEFCRTGTNFTLRELGLGEMTPPHGTLVYRTDIGL PHCYGSMGASSADLEADTVLSPHEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHE NTETGAPLHCSSASSTPIEQSPSPPPANESQRULLNGVAQTPDSDSEEEFVFN SFLVSGSASLGAANDHPSSLQNHPRLRTPPPLPHAHTPNQHHAAASINSLNRGNFT PRSNPSPAPTDHLSGEPAGSAQEPHTAQDNWVLSKIPVETRNLGKLPFLGTWQDN LIEMDIFASRRDGYSDGHFFPKPGGTSPLFTTSPGYPLTSSTVSPPPRLPRST FSRPAFLKKPSKYCNWKAALSAILISATLVILLAYFVAMHLFGLNHLQPMEGQM QMYEITEDASSWPVPTDVSLYPSGGTGLETPDRKGAAEGKPSLFPEDSFIDSGEI DVGRASQKIPPGTFWRSQVFIIDHPVHLKFNVSHTGKAALVGIYGRKGLPPSHTLQDFV ELDDGRRLLTQEARSLGEPQRSRGPVPPSSHETGFIQYLDSGIWHLAFYNDGKESEV VSFLTATIESVDNCPNCGNGDCISGTCFLGFLGPDCCGRASCPFLCGNGQYMKG RCLCHSGWKGAECDVPTNQCIDVACSSHGTCIMGTCICNPGYKGESCEEVDCMDPTCS SRGVCVRGCHCSVGWGTNCETPRATCLDQCSGHGTFELPDTLNCNCDPSPWTGHDCSI EICAADCGGHGVCVGGTCRCEDEGWMGAACDQACHPRCAEHGTCRDGKCECTPGWNGE HCTIEGCPGLCNGNGRCTLDLNGWHCVQCQLGWRGTGCDTSMETGCGDGKNDGDGLVD CMDPDCCLQPLCHVNFCLGSPDPLDIIQETQAPVSQQNLNSFYDRIKFLVGRDSTHS IPGENPFDGGHACVIRGQVMTSDGTPLVGNISFINNPLFGYTI SRQDGSFDLVTNGG ISIIILRFERAPFITQEHTLWLPWDRFFVMTETIVMRHEENEIPSRDLSNFARPNPVVSP SPLTSSFASSCAEKGPIVPEIQALQEEI VIAGCKMRLSYLSRTPGYKSVVRISLTHPT IPFNLMKVHLMVAEGRLFKRWFAAAPDLSYFIWDKTDVYNQKVFGLSEAFVSVGYE YESCPDLILWEKTAVLQGYEIDASKLGGWSLDKHALNIQSGLHKNGENQFVSQQ PPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDENYIRRIFFS GNVNTILEMSHSPAKEYYLATDPMGSAVFLSDTNSRRVFKVSTTVVXDLVKNSEIIVA GTGDQCLPDDTRCDGGKATEATLTNPRGITVDKFGLIYFVDGTMIIRDQNGILST LIGSNDLT SARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYLVNDNNVLIQISENHQV RIVASRPHMHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQ VTSGEISLVAGAPSGCDCKNDANCDFCSDG DYAKDAKLNTPSSLVACDAGELYVAD LGNIRIRFIRKNKPLNTQNMYESPIDQELVYFDTSKGHL YTSQSLFGDLYNFTY TGDGDITHITDNNGMVNVRRDSTGMPLWLVPDQGVYVMTGMTNSALRFSVTTQGH E AMTYSGNSGLLTKSNENGWTTTFYEYDSFGRLTNVTFTPTQGVSSFRSSTDSSVHVQV ETSSKDDVTITTHLSGSGAFYTL LQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHL LAGTVNPTVGKRNVTLPIDNGLNLVWRQKEQARQVTVFGPRLRVNRNLLSLDFD RVTRTKIYDDHRKFTLRILYDQAGRPSFWSPPSRLNGVNVITYSPGGHIAGIQRIMS

ERMEYDQAGRITSRIFADGKMWSYTYLEKSMVLHLHSORQYIFFEDKNDRLSSVTMPN VARQITLIRSVGYRNIYQPPEGNASVIQDFTEDHLLHTFYLCTGRRVIYKYGKLS KLAETLYDTTKVSFTYDETAGMLKTIVNLQNEGFTCTIRYQIGPLIDRQIRFRFTEEGM VNARFDVNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQII TTAVMTHSKHFDAYGRMKEVQYETFRSLMYWMTVQYDNMGRVVKELKVGVPYANTTRY SYEYDADQLQTVSINDKPLWRSYDNLGNLHLLSPGNSARLTPLRDLDRITRLGD VOYKMDRDSLRQRGDVFYNSAGLLIKAYNRASGWSVRYRYDGLGRVSSKSSHSH HLQFFVADLTNPTKVTTHLYSHSSEITSLYDYLQGLHFLAMELSSGDEFYIACDNIQTPLAVFSGTGLMIKQILYATAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMRDDYDVL AGRWTPDHELWKLRLSSNSIVPFHLYMFKNNNPISNSQDIKCFMTDVNSWLLTFGFQL HNVIPGYPKPDTDAMEPSYELVHTQMKTEWDNSKILGVQCEVQKQLKAFVTLERFD QLYGSTITSCQQAPETKKFASGSIKGVKFKALDKGRVTTDIISVANEDGRRIAAIL NNAHYLENLHFTIDGVDTHYFVKPSPSEGLAILGLSGRRRTLENGVNVTVSQINTML SGRTRRYTDIQLQYRALCNTRYGTVDDEKVRVLELARQRAVRQAWAREQQORLREGE EGLRAWTDGEKQQVLNTGRVQGVDFGFFVTSVEQYPELSDSANNIHFMRQSEMGR 1813. .1881 /gene="Doc4" /note="putative; transmembrane-region site" ORIGIN Alignment Scores: Pred. No.: 2,32e-36 Length: 9722 Score: 540.50 Matches: 88 Percent Similarity: 88.12% Conservative: 1 Best Local Similarity: 87.13% Mismatches: 3 Query Match: 87.60% Indels: 9 DB: 10 Gaps: 1 US-10-029-020-14_COPY_750_850 (1-101) x AF059485 (1-9722) QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20 Db 3046 TGGATGGGGCCGCATGCGACCAACGGGCTGCCACCCACGCTGTGCAGAACACGGGACC 3105 QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40 Db 3106 TGCCGGGACGGCAAGTGCAGATGCACCCCGGCTGGAATGGAGAGCACTGCACCATC--- 3162 QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60 Db 3163 -----GAGGGCTGTCTGGCTTGTGCAATGGAATGGCAGA 3198 QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80 Db 3199 TGTACCTGGACCTGAATGGGTGGCACTGTGTCTGCCAGCTGGCTGGCGAGGACTGGC 3258 QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100 Db 3259 TGCACACATCCATGGAAACGGGCTGTGGAGACGGCAAGGACCAACGAGATGGCTTG 3318 QY 101 Val 101 Db 3319 GTG 3321 RESULT 7 AX662357 LOCUS AX662357 8473 bp DNA linear PAT 22-MAR-2003 DEFINITION Sequence 39 from Patent WO02062999. ACCESSION AX662357 VERSION AX662357.1 GI:29163218 KEYWORDS Homo sapiens (human) SOURCE Homo sapiens ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 SpyteK,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X., Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K., Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,B.A., Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F., Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G., Millet,I. and Macdougall,J.R. Proteins and nucleic acids encoding same TITLE

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|:::|||||
Db 2326 TGAAGGATGGCAAGTGAATGCAGCCAGGCTGGAATGGAGAGCACTGCACCTATCGCT 2385
|:::|||||
QY 41 HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
|:::|||||
Db 2386 CACTATTGGATAAGATAGTTAAAGACAAGATAGGATATAAAGAGGGTTGTCTCTGCTCTG 2445
|:::|||||
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
|:::|||||
Db 2446 TGAACAGCAATGGAAGATGTACCTGGACCAAAATGGCGGACATTGTGTGTGCCAGCCT 2505
|:::|||||
QY 75 GlyTyrArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerIlysAsp 94
|:::|||||
Db 2506 GGATGGAGAGGAGCAGGCTGTGACGTAGCCATGGAGACTCTTTGCACAGATACCAAGGAC 2565
|:::|||||
QY 95 AsnAspGlyAspGlyLeuVal 101
|:::|||||
Db 2566 AATGAAGGGATGGACTCATT 2586
|:::|||||

RESULT 10
AB026980
LOCUS AB026980 9264 bp mRNA linear VRT 05-NOV-1999
DEFINITION Danio rerio mRNA for ten-m4, complete cds.
ACCESSION AB026980
VERSION AB026980.1 GI:5307784
KEYWORDS ten-m4.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (sites)
Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.
Compartmentalized expression of zebrafish ten-m3 and ten-m4,
homologues of the Drosophila ten(m)/odd Oz gene, in the central
nervous system
Mech. Dev. 87 (1-2), 223-227 (1999)
99425191
PUBMED 10495292
REFERENCE 2 (bases 1 to 9264)
AUTHORS Mieda,M.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute,
RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hirosawa,
Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp,
Tel:81-48-467-9713, Fax:81-48-467-9714)
COMMENT Sequence updated (29-Jun-1999).
FEATURES
source
1..9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1..9264
/gene="ten-m4"
323..8797
/gene="ten-m4"
/note="similar to Drosophila melanogaster tenm/odz and Mus
musculus Doc4"
/codon_start=1
/product="ten-m4"
/protein_id="BAA81893.1"
/db_xref="GI:5307785"

translation="MEVKRRPYRSLTSRRDTERRYTSSADSEDGKINPKYSSES
LKAFDQDSRLAYGSRVKDLVHHEADEFSRQDPFSLRDMAFGPVPPHMGAYRTEMGL
PHRDYSVSADADTETDGI MSPAHVRLWGRNNTKSGRSSCLFSRANSLTLDTEH
ENTENGPPPLHCSSASSSPVDSPPYPPSHAANQSGRLLGNSGAQAGRDSSEDFGN
SFLVKTGSGNVCTPAAATANEGSFQNHSLRTRTLPPLFHSHPSQHHTASIGSLRSN
YTORNSPAPTSSAPNEGPTSAQDSSSAQDNWLLNSNVPLETRNIAKOTFLETLD
NFIEMDILATARRDGAYTDGHFLFKPGGTSPLYCTTSPGYPLTSSVTSVSPPLPRN
TFSRPAFSLKKPYKHCNWKCAALSAILISVTLVFLAYFIAMHLFGLNWHLPVQRQI

YQLTEDNTSGLHLPTDLGLPLPLNGTGLEFPDRGRSDDGKLDGFFPEDSFIDMGEIDVG
RKVAQLIPGIFWRSQVFIIDHPMYLKFNVSLSKDALVGIYGRRLPPSHQTFDFVELL
DGRRLLSQGLPDGPPFPAQQRSLVPIITSHDTGCIQYMDSGIWHLAVNDGKYTEQV
SELTTAIDSIDDCPSNCFNGDCVSGNCHCFPGFRGPDCSRASCPLVCSGNGQYLKGR
CMCHSGWKGSECDVPTNQCIDITCSGHGTCTVGTICINPSYKGENCEEVDCLDPTCSG
RGVCVRGECHCFVWGPGCESPRASCMEQCSGHGSFLADTNTCNDHNWTFHDCSTE
LCAADCGHGICVAGSCRCDEGWMGTGCEQRACHPRCSEHGTCKDKCECSPGWNGEH
CTIEGCPGLCNGNRCRTLNNNGWYCVCLGWRGACDTSMETACSDGKDGDGLTDC
MDPCCLOASCHTTSCLVSGSPDLIIQETQISSLSLTLSQFYQRIHFLVGRDSTHVI
PDVNPFGIHACVIRGQVVTSDGTPLVGNISFINPKPAYGYTITRQDGSFDLVSNGGV
AIGLRFERAPFITQEHTLWLPWGRFFVMDTIIVMRHEVNDIPSCDLSSTRPMPILVPA
PLTAFAGTCPERGIVVPEIQTLQEEVRIPGTDMRLGYSRTSGYKSLIRITLTSTI
PFSLMKVHLMVAVEGLFRKWFSAAPNLSYDFVWDKTDVYSQKVGLSEAFVSVGFY
ESCPDLIIWEKRTAVLQGYETTASNLGGSVDKHALNIQSGILHKGNGENIFISQQP
PVGISIMNGRRRSISPCPSCNGLADGNKLLAPALACGSGSLYVGFYVRRIRFTTG
NVTSVLELSNSPAHHYYLATSPVSWLYLSDTSSRKVFVKSLYAVKDVAKNLELVAG
TGDQCLPYDETRCGDGKAVEATLNPRTITVDPKGYIFFVDGTMIRRIDQNGIISTL
LGENDLTSARPLSCDSVMDISQVRLWPTDLAVSPMDNSLYLDNNVVLQISENHQVR
IVAGRPMHCQVPLGDHFLVSKTIAHATLESANALAVSHGSLYIAESDEKKINRVQV
STNGEISLLAGAPSGCDCKNDANCDCYSGDDGYAKDAKLANPSSLAVSPDGELFIADL
GNIRYVRNKAFLNPLNMYEISSPIDDELYLFDVNASHVFTQSLTGDLYLNFYTS
GEGDLSITDKNKRVSIIRSDTGLPLWLMGPDGQTFWFTMGTTNNALKSVAAGQGEIA
VMTYHSGSGLLATKSNEDGWSIFYEDNYGRLTNVYPTGRVSSYRTSDSTVRVQTE
GSNKEDITVTNLSASGTFYTLMDQDVKNYSYIIGLGSRLVLANGMEVSLHTEPHLL
SGTNPTISKRNVTLPIDNGLNLVWRQRKEQARQVTVVGRRLRVHNRNLLSMDFDR
VTRTEKYDDHRKFTLRHIDHAGRPTLWAPSSRLGNVNTYSPGGHIAGIORGTMSV
RMEYDQNGRITSKIFADGKSWSYTYLEKSMVLLYSQRYIFEEDKNDRLLSSVTMPNV
ARQTLTTRSIGYRNTYRPPENATVLQDYSDDLQTIHQGTGRRVIYKYGKLSR
LLEILYDTRIAFSYDESAGMLKTVLGLOSEGFACTIRYQIGPLIDRQIFRFSEGMV
NARFDYNDNSFRVTSMAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIIT
TAVMTHTKHFDAYGRVKEVYEIIFRSLMYMMVQFDNMGRRVAKELKVGPYANTTRYA
YEXDADGQLQVVSINDKPLWYSYDLNGLNHLHLLSPENSARLTPLRYDIRDRITRLGDV
QYRLDEDGFLRQGNDFFEYSAGLLVKTYNKVGNTTKYRYDGLGRRVSSRSRTQGH
LOFFYADLSSPTRVTHMYNHSSEITSLYDYLQGLHFAEMLSSGDEFYVACDNIETPL
AVFSGAGLMIKQIILHTAFGEVYLDNSPFLVQIYQGGLYEPLTKLVHMGRRDYDVL
GRWTPDHDIRKRLNSDNIVPFLYMFKNPNPLSNSQETKCYMTDVNSWLVTFFGQLY
NVIPGYRKPVTDAMEPSYELVHTQIKTQEWDSKTVLQCEVQVQRLKSFVRLERFGQ
IYSASDSGCPPTPLHTLFATGSLFGKGVKVAIREGRVEADIIISLANEDGRRIAAVLD
KASYLQDLHFTIAGLDTHYFVKSGLVEGDLSLGTMVGORTLETGVNTVTSQVNMVLG
GRSRRITDIQMOYGTLSLVRYGSSVDEKRVLELARQAVATAWAHERHLRQGEE
GSAWTGDERQQLLSGRVQGYEGFYIVSVDQFPELTNNINNVHFWRQTEMGR"

ORIGIN

Alignment Scores:
Pred. No.: 3.16e-33 Length: 9264
Score: 503.50 Matches: 81
Percent Similarity: 85.00% Conservative: 4
Best Local Similarity: 81.00% Mismatches: 6
Query Match: 81.60% Indels: 9
DB: 5 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x AB026980 (1-9264)

Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|:::|||||
Db 2783 TGGATGGGCACCGGGTGGCGAGCAACGGGCTTGTACACCTCGCTGCAGCAACACGGGACC 2842
|:::|||||
Qy 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|:::|||||
Db 2843 TGAAGGACGGAAAGTGCGRATGCAGCCCGGGATGGAATGGAGAACACTGCATATC--- 2899
|:::|||||
Qy 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|:::|||||
Db 2900 -----GAGGGCTGCCCTGTGTGCAATGGCAATGGCAGA 2935
|:::|||||
Qy 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|:::|||||
Db 2936 TGCACGCTGGGTAATAATGGCTGGTACTCGCTTTGTCAACTGGCTGGCGGGGAGCAGGC 2995
|:::|||||
Qy 81 CysAspThrSerMetGluThrAlaCysGlyAspSerIlysAspAsnAspGlyAspGlyLeu 100
|:::|||||
Db 2996 TGCACACATCAATGGAGACGGCCTGCAGTGACGSCAAGGACCAATGATGGAGATGGGTG 3055
|:::|||||

RESULT 11
AK125869
LOCUS
DEFINITION
Homo sapiens cDNA FLJ43881 fis, clone TEST14009028, highly similar to Mus musculus odd Oz/ten-m homolog 3 (Drosophila) (Odz3).
ACCESSION
AK125869
VERSION
AK125869.1 GI:34532128
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3394)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES
source
1. .3394
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TEST14009028"
/tissue_type="testis"
/clone_lib="TEST14"
/note="Cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Pred. No.: 6.2e-32 Length: 3394
Score: 483.50 Matches: 78
Percent Similarity: 82.18% Conservative: 5
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 78.36% Indels: 9
DB: 9 Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x AK125869 (1-3394)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 491 TGGACGGGCCAGCCTGTATCAGAGAGCCTGCCACCCCGCTGTGCCGAGCACGGGACC 550
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 551 TGCAGGATGGCAAGTGTGATGCAGCCAGGCTGGATGGAGAGCACTGCACATATC--- 607
QY 41 HisTyrLeuAspArgValIleValIleValIleValIleValIleValIleValIleVal 60
Db 608 -----CAGGGTTGCTCCTGCTGTGTGTCAGCCCTGGATGGAGAGCAATGGAAGA 643
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 644 TGTACCTGGACCAAAATGGCTGGCATTTGTGTGTGCCAGCCTGGATGGAGAGGAGGAGGC 703

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 704 TGTGACGTAGCCATGGAGACTCTTTGCACAGATAGCAAGGACAATGAAGGATGGACTC 763
QY 101 Val 101
Db 764 ATT 766
RESULT 12
AF195418
LOCUS
DEFINITION
Mus musculus ODZ3 (Odz3) mRNA, partial cds.
ACCESSION
AF195418
VERSION
AF195418.1 GI:6760368
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 7816)
Ben-Zur,T., Feige,E., Motro,B. and Wides,R.
The mammalian Odz gene family: Homologs of a Drosophila pair rule gene with expression implying distinct yet overlapping developmental roles
Dev. Biol. 117, 107-120 (2000)
2 (bases 1 to 7816)
Ben-Zur,T., Motro,B. and Wides,R.
Direct Submission
Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan University, Ramat-Gan, Israel
Location/Qualifiers
1. .7816
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
2. .7816
/gene="Odz3"
/note="synonym: Ten-m3"
3. .7042
/gene="Odz3"
/note="similar to the Drosophila melanogaster pair rule odd Oz/Ten-m gene sequence"
4. .7042
/codon_start=2
5. .7042
/product="ODZ3"
6. .7042
/protein_id="AAF28316.1"
7. .7042
/db_xref="GI:6760369"
8. .7042
/translation="NSDTPVTNTVSLPSGDNKGLGGFTHEHNTIDSGELDIGRAIOE VPPGIFWRSQLFIDQPQFLKFNISLQKDALIGVYGRKGLPPSHTQXDFVELLDGSRLLI AREQNLVESERAGRARQARSVLSHEAGFIQLDSGIWHLAFYNDGKNPEQVSNFTIVIE SVVECPRNCHNGECVSGTCHCFPGFLGDCSRAACPVLCSGNGQYSGKRCCLCFSGWK GTECDVPTTCIDIDPQCGGRGICIMSGCACNSGYKGENCEADCLDPCGSHGVCIHGE CHCNPGWGSNCEILKTMCADQCSGHGYLQESGCTCDPNWTGPDCSNEICSDVDCGS HGVCMGSCRCCEEGWTGPACNQACHPCRAEHGTCCKGCECSQGWNGEHCETIEGCPG LCNSNGRCTLDQNGWHCVCPQGWGAGCDVAMETLCTDSKONEGDLIDCMPDCCCLQ SLASVIRGQVLTADGTPLIGNVSVFLHYSEGYTITRQDGMFDLVANGASLTLLVFER SPFLTQYHTVWIPWNVYVMDTLVMKKEENDIPSCDLSGFVRPSPIVSSPLSTFFRS SPEDSPIIPETQVLHEETTIPGTDLSVLSRAAGYKSVLKITMTQAVIPFNLMKVH LMVAVVGRFLFOKWFPPASPNLAYTTFIWDKTDAYNQVGLSEAVSVSVGYEYESCLDLTL WEKRTAVLQGYELDASNMGGWTLDKHVLVDQNGILYKNGENQFISQPPVSVSSIMG NRRRSISCSGNCQADGNKLLAPVALACIDGSLYVGFNFYVRRIFPSGNVTSVLEL SSNPAHRYLLATDPTVDLYSDTNTTTRIRYRPKSLTGAKDLTKNAEVVAGTGEQCLPF DEARCGDGGKAVEATLMSPKGMADKNGLIYFVDGTMIRKVDQNGIISTLLGSNDLTS ARPLTCTSMHISQVRLEWPTDLAINPMDNSIYVLDNNVLTQITENRQVRIAAGRPMH CQVPGEYVPGKHAVQTLESATAIAVSGVLYITETDEKINIRIQVTTDGEISLV AGIPSECDKNDANDCYQSGDGYAKDALKNAPSLSAASPDGTLIADLGNIRIRAVS KNPPLNSMNFVEVASTDQELIYFDINGHTQVTSVLTGDTLYLNFYSNDNDVTAVT DSNGTNLTIRIRDRNRMPPVRVSPDNQVITLITGNGCLSKMTAQGLELVLFYHGNSSG LLATKSDETGTWTFDYDSEGLTNTVPTGVTNLTNLTGDMKAITVDIESSSREEDVS ITSNLSSIDSFYTMVQDQLRNSYQIGDGLRIFYASGLDSHYQTEPHVLACTANPTV AKRMTLPGENGQNLVEWRFRKEQAQGVNVFGRKLVNGRNLLSVDFDRTTTEKIY DDHRKFLRLRIAYDTSGHPTLWLPSSKLMVNVVTSYSTGQIASIQRTTSEKVDYDSQG

GFTISRQDGSFDLVAIGGISVILIFDRSPFLPEKRTLWLPWNQFIIVVEKVTMQRVVSD
PPSCDISNFISPNPIVLPSPLTSFGGSCPERGTIVPELQVVQEEIPIPSFVRLSYLS
SRTPGYKTLRLILLTSTIPVGMIKVHLTVAVEGRLTQKWPAAINLVYTFAMNKTDI
YQQVWGLAEALVSVGEYETCPDFILWEQRTVVVLQGFEMDASNLGDWSLNKHHILNP
QSGIIHKNGENMFISQQPPVISTIMNGHQRSVACTNCNGPAHNKLFAPVALASGP
DGSVYVGFNFVRRIFPSGNSVSIIELESTSPAHKYILAMPVSESLYLSDTNTRKVIYK
LKSIVETKDLKSNFEVVAGTDQCLPFDOQSHCGDGRASEASINSPRGITVDRHGFIY
FVDGTMIRKIDENAVITTVIGSNGLTSTQPLSCDSGMDITQVRLIEWPTDLAVNPMDNS
LYVLDNNIVLOISENRRVRIAGRPVHCQVPGIDHFLVSKVAIHSTLESARAI SVSHS
GLLFAETDERKVNRIQQVTTNGEIIYIAGAPTDCKIDPNCDCFSGDGGYAKDAKM
KAPSSLAVSPDGTLYVADLGNVRIRTIISRNQAHLMNMIYEIASPADQELYQFTVNGT
HLHTLNLITRDYVYNFTYNSGDLGAIITSSNGNSVHIRRDAGGMPLWLVPGGQVYWL
TISSNGVLKRVSAAQYNPALMTYPGNTGLLATKSNENGWTVVYEYDPEGHLTNATFPT
GEVSSFHSDLKLTKEVLDTSNRENVLMSTNLTATSTIYILKQENTQSTYRVNPDGSL
RVTFASGMEIGLSSEPHILAGAVNPTLGKCNISLPGEHANLIEWRQKEQNKGNVSA
FERRLRAHNRNLLSIDFDHITRTGKIYDDHRKFTRLILYDQTRPILWSPVSRNEVN
ITYSPSGLVTFIQRTWNEKMEYDQSGKIIISRTWADGKIWSYTYLEKSVMLLHSQRR
YIFEYDQSDCLLSVTMPMSVVRHSLQTMLSVGYRYNIYTPPDSSTSFIDYSRDGRLLQ
TLHLGTGRRVLYKYTKQARLSEVLYDTTQVTLTYESSGVIKTIHLMHDGFICTIRYR
QTGPLIGRQIFRFSEEGLVNARFDYSNNFRVTSMAVINETPLPIDLRYVDVSGRT
EQFGKFSVINYDLNQVITTTVMKHTKIFSANGQVIEVQYEILKAIAYWMTIQYDINVGR
HGNMCIRVGVVDANITRYFYEYDADGQLQRTVSVNDKTQWRYSYDLNGDINLLSHGKSAR
LTPRLRYDLRDRIITRLGEIQYKMWDEDFGLRQRGNDIPEYNSGLLQKAYNKASGWTVOY
YYDGLGRRVASKSLGQHLQFFVDATANPIRVTHLYNHTSSEITSLYYDLQGLIAME
LSSGEEXYVACDNTGTPLAVFSSRGQVIKEILYTPYGDIYHDTYPDFQVIIGFHGGLY
DFLTKLVLHGLQRDYDVVAGRWTTAYHHIWKQLNLPLPKPNLYSFENNYPVGKIQDVAK
YTTDIRSWLELFGQLHNVLPFPKPELENLELTYELLRLQTKTQEWDPGKTIILGIQC
ELQKQLRNFISLDQLPMTPRYNDGRGLEGGKQPRFAAAPSVPFGKGIKFAIKDGIVTAD
IIGVANEDSRRLLAAILNNAHYLENLHFTIEGRDTHYFIKLGSLSEEDLVLIGNTCGRRRI
LENGVNVTVSQMSTSLNGRTRRFADIQLOHGALCFNIRYGTTVEEKKNHVLEIARQRA
VAQAWTKEQRRRLQEGEERAWTEGEKQQLLSTGRVQGYDYFVLSVEQYLELSDSAN
NIHFMRQSEIGRR"

ORIGIN

Alignment Scores:

Pred. No.:	6.02e-30	Length:	8297
Score:	464.50	Matches:	71
Percent Similarity:	82.00%	Conservative:	11
Best Local Similarity:	71.00%	Mismatches:	17
Query Match:	75.28%	Indels:	1
DB:	9	Gaps:	1

US-10-029-020-14_COPY_750_850 (1-101) x AF100772 (1-8297)

QY	1	TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr	20
Db	2210	TGGGTAGGACCAACATGTGAGGAACGCTCCTGTCAATTCATTTACTGAGCATGGCCAA	2269
QY	21	CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla	40
Db	2270	TGCAAGATGGAAAATGTGAGTGTAGCCCTGGATGGAGGCGGACCACTGCACAATTGCT	2329
QY	41	HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg	60
Db	2330	CACTACTTAGAT---GCTGTCCGAGATGGCTGCCAGGGCTCTGCTTTGGAATGGACGA	2386
QY	61	CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly	80
Db	2387	TGTACCCCTGGATCAAAATGGTTGGCACTGTGTGTCTCAGTGGGTTGGAGTGGGACAGGC	2446
QY	81	CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu	100
Db	2447	TGCAATGTTGTCAATGGAATGCTTTGTGGAGATAACTTGGACAATGATGGAGATGGTTTA	2506

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 2670.71 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400
Perfect score: 782
Sequence: 1 YIRRIFFSGNVTNILELRNK.....LLGSNDLTSARPLSCDSVMD 151

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

Result No.	Score	Query Match %	Length	DB ID	Description
1	782	100.0	8354	6	AX556500 Sequence
2	767	98.1	8585	10	AB025413 Mus muscu
3	759.5	97.1	8438	6	AX675551 Sequence
4	727.5	93.0	8645	6	AX600210 Sequence
5	712.5	91.1	9722	10	AF059485 Mus muscu
6	625	79.9	8409	5	GGA279031 Gallus ga
7	625	79.9	8409	6	AX250068 Sequence
8	619	79.2	3394	9	AK125869 Homo sapi
9	619	79.2	5309	9	AB040888 Homo sapi
10	619	79.2	8473	6	AX662357 Sequence
11	619	79.2	8487	6	AX662359 Sequence
12	619	79.2	8645	6	AX662355 Sequence
13	619	79.2	8675	6	AX662353 Sequence
14	618	79.0	5804	10	AK122513 Mus muscu
15	618	79.0	8964	10	AB025412 Mus muscu
16	613	78.4	6560	6	AX250065 Sequence
17	613	78.4	7781	9	AB032953 Homo sapi
18	613	78.4	9729	6	AX250013 Sequence
19	613	78.4	9826	6	AX250008 Sequence
20	610	78.0	8689	6	AX250067 Sequence
21	610	78.0	8689	10	AF086607 Rattus no
22	610	78.0	8797	6	AX250063 Sequence
23	610	78.0	8797	6	AX250066 Sequence
24	610	78.0	8797	10	AB025411 Mus muscu
25	608.5	77.8	9264	5	AB026980 Danio rer
26	602.5	77.0	8575	6	AX921803 Sequence
27	588	75.2	8373	10	AB025410 Mus muscu
28	585	74.8	8118	5	GGA238613 Gallus ga
29	571.5	73.1	7713	9	HSM808325 Homo sapi
30	563.5	72.1	7816	10	AF195418 Mus muscu
31	555.5	71.0	8816	5	AB026979 Danio rer
32	534.5	68.4	8297	9	AF100772 Homo sapi
33	413	52.8	112872	2	AP000716 Homo sapi
34	413	52.8	165812	2	AP001141 Homo sapi
35	413	52.8	170156	2	AP002412 Homo sapi
36	413	52.8	186084	9	AP002768 Homo sapi
37	406	51.9	245722	2	AC120288 Rattus no
38	406	51.9	257693	2	AC103323 Rattus no
39	405	51.8	226256	10	AC073599 Mus muscu
40	388	49.6	179685	5	BX001057 Zebrafish
41	347	44.4	321049	2	AC137221 Rattus no
42	344	44.0	193544	5	AL807754 Zebrafish
43	308.5	39.5	60511	9	HS618F1 Human DNA
44	307	39.3	146071	10	AL929061 Mouse DNA
45	307	39.3	172465	2	BX324192 Mus muscu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

AX556500 LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,
Smithson,G., Li,L. and Ji,W.
TITLE Polypetides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1. .8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 4.41e-79 Length: 8354
Pred. No.: 782.00 Matches: 151
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6
US-10-029-020-14_COPY_1250_1400 (1-151) x AX556500 (1-8354)
Qy 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTAGAGGATCTTCCCTCTGGAATGTCAACCAATCCTAGAGTGAGGAATAAA 3841
Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3842 GATTTCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGT 3901
Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGGCCGTCTTCTCTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTG 3961
Qy 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGTTGCGGGACAGGTGACCATGCCTC 4021
Qy 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTGCGGGGATGGTGGGAGGCCACAGAAGCCACACTACCAAT 4081
Qy 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCCAGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGATC 4141
Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGCATCGATCAGAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATCA 4201
Qy 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTCAGCTGTGATCTGTCTCATGGAT 4234
RESULT 2
AB025413
LOCUS AB025413 8585 bp mRNA linear ROD 08-MAY-1999
DEFINITION Mus musculus mRNA for Ten-m4, complete cds.
ACCESSION AB025413
VERSION AB025413.1 GI:4760781

Ten-m4.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (sites)
Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8585)
Oohashi,T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail:oohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128,
Fax:+81-86-222-7768)
Location/Qualifiers
1. .8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1. .8585
/gene="ten-m4"
188. .8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translation="MDVKERKPYRSLRRRDAERRYTSSSADSEEGKGPQKSYSSSET
LKAYDQDARLAYSRVKMDVPQEAEEFCRTGTNFTRELGLGEMTPPHGTLRYTDIGL
PHCGYSMGASSDADLEADTVLSPHPVRLWGRSTVSRSGRSSCLSSRANSNLTLDTEHE
NTETDHPSSLQNHPRLRTPPPLPHAHTPNQHHASINSINLRGNFTPRSNPSPAPTDH
SLSGEPPAGSAQEPHTAQNWLNSNIPLETRNLGKQPFLGTLQDNLIEMDILSASRH
DGAYSDGHFLFKPGGTSPLFCTTSPGYPLTSTSTVSPPPRPLPRSTFSRPAFNLKPS
KYCNWKAALSAILISATLVILLAYFVAMHLFGLNHLQPMEGQMOMYEITEDTASSW
PVPTDVSLYPSGGTGLETPDRKGKGAAGKPSLPEDSFIDSGEIDVGRRASQKIPP
GTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIYGRKGLPPSHTQDFDVELLDGRRLLTQE
ARSLGPQRQSRGPVPPSSHETGFIQYLDSGIWHLAFYNDGKESEVVSFLTIAIESVD
NCPSNCYNGDCISGTCHCFLGLGPDGCRASCPVLCSGNGQYMKGRCLCHSGWKGA
CDVPTNQCIDVACSSHGTCTIMGTCTCNPGYKGESCEEVDMDFTCSSRGVCVRGECHC
SVGWGTCNCTPRATCLDQCSHGHTFLPDGTGLCNCDPSWTGDCSIEICAADCGGHG
CVGGTCRCEDGWMGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHCHTIAHYLDRVV
KEGCPGLCNGNRCRTLNLNGWHVCVQLGWRGTGCDTSMETCGCGDKNDGDGLVDCMD
PDCLOPLCHVNPLCLGSPDPLDIQETQAPVSNQNLNPFYDRIKFLVGRDSTHSIPG
ENPFDGHACVIRGQVMTSDGTPLVGVNISFINNPLFGYTSIRQDGSFDFLVTNGGISI
ILRFERAPFIQEHTLWLPWDRFFVMEIIVMRHEENIIPSCDLSNEFARNPNVSPSPL
TSFASCAEKGPVPEIQALQEEIVIAGCKMRLSYLSRTPGYKSVLRISLTHPTIFE
NLMKVHLMVAVEGRLFRKWFAAAPDLSYFIWDKTDVYNQKVFGEAFVSVGYEYES
CPDLILWEKRTAVLQYIEIDASKLGSWLDKHALNIIQSGILHKGNGENQFVSQQPPV
IGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFNYIRRIFFSGNV
TNILEMRNKDFRHSHPAHKYIYATDPMGSAVFLSDTNSRRVFKVKSITTVVKDLVKNS
EVVAGTGDQCLPFDDTRCGDGGKATEATLTNPRGITHVDKFGLIYFVDGTMIRRVQNG
IISTLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYLDNNVVLQISE
NHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKN
RIRQVTTSGEISLVAGAPSGCDCKNDANDCDFSGDDGYAKDAKLNTPSSLAVCADGEL
YVADLGNIRIRIRKNKPFNTQNMVELSSPIDQELYLEFDTSGKHLYTQSLPTGDLY
NFTYTGDDGDIHTIDNNGNMNVNRDSTGMPLWLVPDQVQVYVWTMGTSALRSVTQ
GHELAMTYHGNSGLLATKSNENGWTFYEYDSFGRLTNTVFTPTGQVSSPRSDTSSV
HVQVETSSKDDVTITTNLSASGAFYTLIQDQVRNSYIIGADGSLRLLLANGMEVALQT
EPHLLAGIVNPTVGKRVNLTLPIDNGLNLVWRQRKEAQGVTVFGRRLRVHNRNLLS
LDEFVRTRTEKIYDDHRKFTLRILYDQAGRPSPSPSRSLNGVNVITYSPGHHIAGIQR
GIMSERMEYDQAGRITSRIFADGKMWSYTYLEKSMVLHLHSORQYIFEDKNDRLSSV
TMPNVARQTLLETIRSVGYRNIYQPPEGNASVIQDFTGHLHTFYLGTRRVIYKY
GKLSKLAETLYDITKVSFTYDETAGMLKTVNLQNEGFTCTIRYQIIGPLIDRQIFRFT
EEGMVNARFDVNDNSFRVTSMQAVINETPLIDLYRDDVSGKTEQFGKFGVIYDI
NQIITTAVMTHTKHFDAYGRMKVEQYEIFRSLMWMTVQYDNGMRVVKELKVGPPYAN

TTRYSEYDADGQLQTVSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDLDRIT
RLGDVQYKMDDEGFLRQGGDVFEYNSAGLLIKAYNRASGWSVRYRYDGLGRRVSSKS
SHSHLLQFFSYADLTNPTKVTYHLNHSSEITSLYDLDLQGLHFAFAMELSSGDEFYIACDN
IGTPLAVFSGTGLMIKQILYAYGEIYMDTNPNQIIGYHGGLYDPLTKLHVHMRDD
YDVLAGRWTSPDHELWKRLSSNSIVPFHLYMFKNNPISNSQDIKCFWTDVNSWLLTF
GFQLHNVIPGYPKPTDAMEPSYELVHTQMKTBWDNSKSIILGVQEVQKQKAFVTL
ERFDQLYSTITSQOAPETKKFASSGSI FGKGVKFPALKDGRVTTDIIISVANEDGRRI
AAILNNAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGGRRITLNGVNTVTSQI
NTMLSGRTRRYTDIQLRYALCLNTRYGTTVDREKVRVLELARQRAVRQAWAREQORL
REGEGLRAWTDGEKQQLNTGRVQGYDGFVTSVEQYPELSDSANNIHFMRQSEMR
R"

ORIGIN

Alignment Scores:
Pred. No.: 2.41e-77 Length: 8585
Score: 767.00 Matches: 145
Percent Similarity: 99.34% Conservative: 5
Best Local Similarity: 96.03% Mismatches: 1
Query Match: 98.08% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AB025413 (1-8585)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3941 TACATCCGCAGAAATCTTCCCCTCTGGAAATGTCCACCAACATCCTGGAGATGAGAAATAAA 4000
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 4001 GATTTCAGACATAGTCACAGCCCAGCACACAAATACTACCTGGCTACAGACCCCATGAGT 4060
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 4061 GGGGCCGCTTTCCTGTCTGACACCAACAGCCGGCGGTCTTCAAGGTCAAGTCCACCACA 4120
QY 61 ValValLysAspLeuValLysAsnSerGluValAlaGlyThrGlyAspGlnCysLeu 80
Db 4121 GTGGTGAAGGACCTGGTGAAGAACTCCGAGTGCTGAGGACTGGTGACCAAGTGCCTC 4180
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4181 CCCTTTGATGATACCCGCTGCGGAGATGGTGGGAAAGCCACAGAACCGCTCACTAAC 4240
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4241 CCCAGGGGAATTACAGTGGACAAAGTTGGGCTCATTTATTTCGTGGACGGCACCATGATC 4300
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4301 AGACGTGTGTGATCAAAATGGAATCATCTCCACTTTGCTGGGCTCCAATGACCTCACCTCG 4360
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4361 GCCAGGCCCTCAGCTGTGACTCCGTCATGGAG 4393

RESULT 3
AX675551
LOCUS AX675551 8438 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 1 from Patent WO02055704.
ACCESSION AX675551
VERSION AX675551.1 GI:293333552
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Bolidog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J., and
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and

TITLE Stone,D.J.
JOURNAL Proteins, polynucleotides encoding them and methods of using the
same
FEATURES
source Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
Location/Qualifiers
1. .8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 1.72e-76 Length: 8438
Score: 759.50 Matches: 151
Percent Similarity: 96.79% Conservative: 0
Best Local Similarity: 96.79% Mismatches: 0
Query Match: 97.12% Indels: 5
DB: 6 Gaps: 2

US-10-029-020-14_COPY_1250_1400 (1-151) x AX675551 (1-8438)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeu-----Arg 18
Db 3764 TACATTAGAAGGATCTTCCCCTCTGTGAAATGTCCACCAACATCCTAGAGCTGAGGGTCAGA 3823
QY 19 AsnLysAspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPro 38
Db 3824 AATAAAGATTTCCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCC 3883
QY 39 MetSerGlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSer 58
Db 3884 ATGAGTGGGGCGCTTCTCCTTTCTGACAGCAACAGCCGGCGGTCTTTTAAATCAAGTCC 3943
QY 59 ThrValValValLysAspLeuValLysAsnSerGluValAlaGlyThrGlyAspGln 78
Db 3944 ACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCAG 4003
QY 79 CysLeuProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeu 98
Db 4004 TGCCTCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAGCCACAGAACCCACACTC 4063
QY 99 ThrAsnProArg-----GlyIleThrValAspLysPheGlyLeuIleTyrPheVal 115
Db 4064 ACCAATCCAGGGGTCCCCAGGCATTACAGTGGACAAGTTTGGGTGATCTACTTCGTG 4123
QY 116 AspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySer 135
Db 4124 GATGGCACCATGATCAGACGCATCGATCAGATGGGATCATCTCCACCTGCTCGGCTCT 4183
QY 136 AsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMetAsp 151
Db 4184 AATGATCTCACATCAGCCCGGCCACTCAGCTGTGATTCTGTCTGATGGAT 4231

RESULT 4
AX600210
LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S.,
Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,
Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K.,
Khare,R. and Wallia,N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)

FEATURES source Location/Qualifiers 1..8645 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="Incyte ID No: 7488573CB1"

ORIGIN

Alignment Scores: Pred. No.: 8.43e-73 Length: 8645 Score: 727.50 Matches: 144 Percent Similarity: 95.36% Conservative: 0 Best Local Similarity: 95.36% Mismatches: 0 Query Match: 93.03% Indels: 7 Gaps: 1 Db: 6

US-10-029-020-14_COPY_1250_1400 (1-151) x AX600210 (1-8645)

Qy 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20 Db 3852 TACATTAGAAGGATCTTCCCTCTGGAAATGTCACCAACATCCTAGAGCTG----- 3902 Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40 Db 3903 -----AGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGT 3950 Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60 Db 3951 GGGCGCGTCTCTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTG 4010 Qy 61 valVallysAspLeuVallysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80 Db 4011 GTGGTGAAGGACTTGTCAAGAACTCTGAGGTGGTTGCGGGACAGGTGACAGTGCCTC 4070 Qy 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100 Db 4071 CCCTTTGATGACACTCGCTGCGGGATGGTGGGAAGGCCACAGAGCCACACTCACCAAT 4130 Qy 101 ProArgGlyIleThrValAspLysPheGlyLeulleTyrPheValAspGlyThrMetIle 120 Db 4131 CCCAGGGGCATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGATC 4190 Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140 Db 4191 AGACGCATCGATCAGATGGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATCA 4250 Qy 141 AlaArgProLeuSerCysAspSerValMetAsp 151 Db 4251 GCCCGGCCACTCAGCTGTGATCTGTGTCATGGAT 4283

RESULT 5 AF059485 AF059485 9722 bp mRNA linear ROD 15-AUG-1998 LOCUS Mus musculus DOC4 (Doc4) mRNA, complete cds. DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds. ACCESSION AF059485 AF059485.1 GI:3170614 VERSION . KEYWORDS SOURCE Mus musculus (house mouse) ORGANISM Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 9722) Wang,X.Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P., Zinszner,H. and Ron,D. Identification of novel stress-induced genes downstream of chop EMBO J. 17 (13), 3619-3630 (1998) MEDLINE 98315054 PUBMED 9649432 REFERENCE 2 (bases 1 to 9722) Wang,X.-Z. and Ron,D. Direct Submission TITLE Submitted (14-APR-1998) Skirball Institute, New York University Medical Center, 550 First Ave., New York, NY 10016, USA JOURNAL

Location/Qualifiers 1..9722 /organism="Mus musculus" /mol_type="mRNA" /strain="NIH/Swiss" /db_xref="taxon:10090" /cell_line="NIH-3T3" /note="tunicamycin-treated" 1..9722 /gene="Doc4" /note="downstream of CHOP 4" 583..9060 /gene="Doc4" /note="similar to Drosophila melanogaster tenm/odz and human gamma-heretulin; type II transmembrane protein" /codon_start=1 /product="DOC4" /protein_id="AAC31807.1" /db_xref="GI:3170615" /translation="MDVKERKPYRSLRRRDAERYTSSADSEEGKGPQKYSSET LKAYDQARLAYGSRVKDMVPOEAEFCRTGTNFTLRELGLGEMTPPHGTLRTDIGL PHCGYSMGASSDADLEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLLDTEHE NTETGAPLHCSSASSTPIEQSPSPPPPPANESQRRLLGNQVAQPTPDSSEEEFVPN SFLVKSASGLGVAANDHPSSLQNHPRLRTPPPLPHAHTPNQHHASINSLNRGFT PRSNPSPAPTDHSLSGEPAGSAQEPHTAODNWVLSKIPVETRNLGKQPFGLTWQDN LIEMDIFSASRRDGAYSDGHFFFKPGGTSPLFTCTTSCGYPLTSTTVSPPPPLPRST PSRPAFNLKKPSKYNKCAALSAILISATLVLLAYFVAMHLFGLNWHLQPMEGQM QMYEITEDASSWPVPTVSLYPSGGTGLETPDRKKGAAEGKPSLFPEDSFIDSGEI DVGRRASQKIPPGTFWRQVFIIDHPVHLKFNVSGLKAAALVGIYGRKGLPPSHTQLDFV ELLDGRLLLTQEARSLLEGQRQSRGPVPPSSSHETGFIQYLDSGIWHLAFYNDGKESEV VSFLTIAIESVDNCPNSCVNGDCISGTCFGLFGLPDCGRASCPVLCSGNGQYMKG RCLCHSGWKGAECVPTNQCIDVACSSHGTCIMGTCICNPGYKGESCEEVDGMDPTCS SRGVCVRGECHCSVGGTNCETPRATCLDQCSHGTFLPDGLCNDPSTGHDCSI EICAADCGGHGVCVGGTCRCEGDMGAAACQDRAHPRCAEHGTCRDGKCECTPGWNGE HCTIEGCPGLCNGNRCITLNGWHCVQLGWRGTGCDTSMETGCGDKNDGDGLVD CMDPCCLOPLCHVNPCLGSPDLDIIQETQAPVSQQNLNSFYDRKFLVGRDSTHS IPGENPFDGGHACVIRGOVMTSDGTPLVGNISFINNPLFGYIISRQDGSFDLVNNGG ISIIIRFERAPFITQEHTLWLPWDRFFVMETIVMRHEENIEIPSRDLNFAFPNPVPS SPLTSFASCAEKGPVPEIQALQEEIVIAICKMRLSYLSRTPGYKSVVRISLTHPT IPENLMKVHLMVAVEGRLEKFWFAAAPDLISYFIWDTDVYNQKVFGLSEAFVSVGYE YESCPDLILWEKRTAVLQGYEIDASKLGGWSLDKHALNIQSGILHKGNGENQFVSQQ PPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGSGSLYVGFDFNYIRRIFFS GNVNILEMSSHSPAHKYYLATDMSGAVFLSDTSNRRRVFKVKTSTVVKDLVKNSEVVA GTGQCLPFDDTRCGDGGKATEATLTNPRGITVDKFLIYFVDTMIRRDQNGIIST LLGNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYLDNNVVLQISENHQV RIVAGPMDHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKNIRIQ VTTSGEISLVAGAPSGCCDKDANDCFSGDDGYAKDAKLTNPSSLAVCADGELYVAD LGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLEFSGKHLVLTQSLPTGDLYLNFTY TGDGDITHITDNNGMVNRDSTGMPFLMWLPDQVYVWVTMTGNSALRSVTTQCHEL AMTYHNSGLLATKSNENGTTFEYDSFGRLTNVTFPTGQVSSFRSDTSSVHVQV ETSSKDDVTITTHLSGSAFYTLQDQVNSYIGADGSLRLLLANGMEVALQTEPHL LAGTVNPVGNKRVNLTLPIDNGLNLEWRQKEAQGVTVFGPRLRVHNRNLLSLDFD RVTRTEKIYDDHRKFTLRILYDQAGRPSFWSPSSRLNGVNVTSYSPGHIAGIQRGIMS ERMEXDQAGRITSRIFADGKMSYTYLEKSMVLHLSQRYIFEEFKNDRLSSVTMPN VARQTLTIRSVGYRNIYQPPGNSAVIQDFTEDGHLHLLTFLYLGTRRVYKYKGLS KLAETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCTIRYQIQGLIDRQIFRFTEEGM VNARFDVYDNSFRVTSMQAVINETPLPIDLYRDDVSGKTEQFGKFGVIYDINQI1 TTAVMTHSKHFDAYGRMKEVQYETFRSLMYMTVQYDNMGRVVKELKVGPIYANTRY SYEYDADGQLQTVSINDKPLWRSYDNLGNLHLSPGNSARLTPLRLDLDRITRLGD VQYKMDDEGSLRQGGDVFEYNSAGLLIKAYNRASGWSVRYRGLGRRVSSKSSHSH HLQFFYADLTNPTKVTHLYSHSSSEITSLYYDLQGLHFMELSSGDEFYIACDNITGP LAVFSGTGLMKIQLIYTAGYGEIYMDTNPNFQIIIGYHGLYDPLTKLVHMGRRDYDLV AGRWTSPDHELMKRLSSNSIVPFHLYMFKNNPISNSODIKCEMTDVNSWLLTFFGFQL HNVI PGYPKPDIDAMEPSYELVHTQMTQEDMNSKSLGVQCEVQKQKAFVTLERFD QLYGSTTSCQAPETKKFASSGSI FGKGVKFAKLDGRVTTDIISVADGGRRIAAIIL NNAHYLENLHFTIDGVDTHYFVKPGPSEGDLAILGSGRRITLNGVNVTVSQINTML SGRTRRYTDIQLQYRALCLNTRYGTTVDDEKVRVLELARQAVRQAWAREQQRIREGE EGLRAWTDGEKQQVLTGRVQYGDYGFVTSVEQYPELSDSANNIHFMRQSEMGRR" 1813..1881 /gene="Doc4" /note="putative; transmembrane-region site"

misc_feature

ORIGIN

Alignment Scores:

Pred. No.: 5.14e-71 Length: 9722
Score: 712.50 Matches: 138
Percent Similarity: 94.70% Conservative: 5
Best Local Similarity: 91.39% Mismatches: 1
Query Match: 91.11% Indels: 7
DB: 10 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x AF059485 (1-9722)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
|||
Db 4519 TACATCGCAGAAATCTTCCCTCTGGAAATGTCACCAACATCCTGGAGATG----- 4569

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
|||
Db 4570 -----AGTCACAGCCCAGCACACAAATACTACCTGGCTACAGACCCCATGAGT 4617

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
|||
Db 4618 GGGGCCGCTCTCTGCTGTGACACCAACAGCCGGGGGTCTTCAAGGTCAAGTCCACCACA 4677

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
|||
Db 4678 GTGGTGAAGGACCTGGTGAAGAACTCCGAGGTGGTAGCAGGGACTGGTGACCATGCCTC 4737

QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
|||
Db 4738 CCCTTTGATACATCCCGCTGCGGAGATGGTGGGAAGCCACAGAGCACGCTCACTAAC 4797

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
|||
Db 4798 CCCAGGGAATTACAGTGGACAAAGTTGGGCTCATTTATTCGTGGACGGCACCATGATC 4857

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
|||
Db 4858 AGACGTGTGTGATCAAAATGGAATCATCTCCACTTTGCTGGGCTCCAATGACCTCACCTCG 4917

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
|||
Db 4918 GCCAGGCCCTCAGCTGTGACTCCGCTCATGGAG 4950

RESULT 6
GGA279031 8409 bp mRNA linear VRT 18-MAR-2001
LOCUS
DEFINITION
Gallus gallus partial mRNA for teneurin-2 (TEN2 gene), long splice variant.
ACCESSION
AJ279031
VERSION
AJ279031.1 GI:10241573
KEYWORDS
ten2 gene; teneurin-2.
SOURCE
Gallus gallus (chicken)
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
1
AUTHORS
Tucker,R.P., Chiquet-Ehrismann,R., Chevron,M.P., Martin,D.,
Hall,R.J. and Rubin,B.P.
TITLE
Teneurin-2 is expressed in tissues that regulate limb and somite
pattern formation and is induced in vitro and in situ by FGF8
JOURNAL
Dev. Dyn. 220 (1), 27-39 (2001)
MEDLINE
20581705
PUBMED
11146505
REFERENCE
2 (bases 1 to 8409)
AUTHORS
Chiquet-Ehrismann,R.
TITLE
Direct Submision
JOURNAL
Submitted (13-SEP-2000) Chiquet-Ehrismann R., Cell Biology,
Friedrich Miescher Institute, Maulbeerstrasse 66, CH-4058 Basel,
SWITZERLAND

FEATURES

source
1..8409
/organism="Gallus gallus"
/mol_type="mRNA"

/db_xref="taxon:9031"
/tissue_lib="11 day and 14 day embryonic brain; 11 day
embryonic retina"
1..8409
/gene="TEN2"
<1..>8406
/gene="TEN2"
/function="pattern formation and neuronal development"
/codon_start=1
/evidence=experimental
/product="teneurin-2"
/protein_id="CAC09416.1"
/db_xref="GI:10241574"

/translation="MDIKDRRHRSLTRGCGKECRYTSSSLDSEDCRVPAQKSYSSSE
TLKAYGHDTRMHYGNRVSDLVHRESDEFPRQGTNFTLAELGICEPSPHRSGYCSDIGI
LHQYSLSTGSDADSDTEGGMSPHAIIRLWGRGKSSRSSGLSSRENSALTITDSNE
NKSDENDFHTLSEKLKDRQTSWQOLAETKNSLIRRPPTPTSSSSLLPSAQLPSSHN
PPPVSCQMPLLDSNTSHQIMDINDEEFPNSYLLRACSPQOASSGSPSNHHSQSTL
RPPLPPHNHSLSHHSSANSLNRSNTNRNQIHAPAPAPNDLATTPESVQLQDSWV
LNSNVPLETRHFLFKTSSGTTPLFSSSSPGYPLTSGTVVTPPRLPRNTFSRNAFKL
KKPSKYCSWKCAALSAIAAAVLAAILLAYFIAMHLGLNWQLQPADGTFNSGLRPGA
AGAEAGAAAPPAGRGPWVTRNSSIDSGETEVRKVTQVPPGVFWRSQIHISQPOFLK
FNISLGDALFGVYIRRGLPSPSHAQYDEMERLDGKEKWSVESPRRRSIQTLVQNEA
VFVQYLDVGLWHLAFYNDGDKKEVVSFSTVILDSVQDCPRNCHNGECVSGVCHCFPG
FHGADCAKAAACPVLCSGNGQYSKGTCLCYSGWKGPCECDVPISQCIDPSCGGHGCIEG
NCVCSIGYKGENCEEVDCLDPTCSNHGVCVNGECLGSPGWGINCELPAQCPDQCSG
HGTYLSDTGLCSCDPNMWGPDCSVEVCVDCGTHGVGICGACRCEEGWTGVACDQVVC
HPRCTEHGCTKDCKECEGWNGEHCTIGRQTTGTETDCPDLCLNGNGRCTLGQNSWQ
CVCQTGWRGPGCNVAMETSACDNKDNEGDLVDCLVPCCLQSTCONSLLCGRSDPL
DIIQQSHSGSPAVKSFYDRIKLLVGKDSHIIIPGENFPNSSLSVLIRGOVVTDTGTP
VGVNVSVFKYPKYGYTITRQDGMFDLVANGSSSLTLHFERAPFMSQERTVWLPWNSFY
AMDTLVMKTEENSIPSCDLSGFVRPDPVIISSPLSTFFSDAPGRNPPIVETQVLHEEI
EVPGSSIKLIYLSRRTAGYKLLKIIMTQSLVPLNLIKVHLMVAVEGHLFOKSFLLASP
NLAYTFIWDKTDAYGQVYGLSDAVVSVGFYETPCPSLILWEKRTALLQGFELDPNLL
GGWLDKHHVNLNVKSGILHKNGENOFLTQQPAVITSIMGNRRRSISCPSCNGLAEG
NKLAPVALAVGIDGSLFVGDFNYIRRIFFSRNVTSILELRNKEFKHNSNPAHKYYLA
VDPVSGSLYVSDTNSRRYIKVKSLTGTKOLAGSEVVAGTGEQCLPFDEARCGDGGKA
VDATLMSPRGIAVDKYGLMYFVDATMIRKVDQNGIISTLLGSNDLTAVRPLSCDSSMD
VSQVRLEWPTDLAVDPMDNSLYLVENNVLIRITENHQVSI IAGRPMHCQVPGIDYLSL
KLAIHSALESASAIASHTGVLYISETDEKINRLRQVTNGEICLLAGASDCDCKN
DVNCNCSGDDGVATDAILNSPSSLAVPDGTIYIADLGNIRIRAVSKNRPILNSFNQ
YEASPGEQELYVFNADGIHQYTLSTLVGTGEYLNFTYSSDNDVTEVMDNSGNSLKVRR
DASGMRHLMPDNQIVTLAVGTNGGLKLVSTQTLGLMTYNGNSGLLATKSDETGW
TTFYDHEGRLTNVTRPTGVVTSLHREMEKSIITIDIENSNRDDDDVTVITLSSVEAS
YTVVQDQVRNSYQLCNNGTLRVMYANGMSISFHEPHVLAGTVPTIGRCNLSLPMEN
GLNSIEWRLRKEQIKGKVTVFGKLRVHGRNLLSIDYDRNIRTEKIYDDHRRKFTLRII
YDQLGRPFELWLPSSGLAAVNVSYFFENGLAQQRGAMSERDIDKQGRIISRMPADGK
VMSYTYLEKSMVLLQLSQORQYIFEYDSSDRDLHAVTMPSVARHSMSTHTSVGIIRNIYN
PPESNASVIFDYDDGRILKTSFLGTGRQVFKYKGLSKLSEIYVDSVAVTFGYDETT
GVLKMVNLQSGGFSCTIRYRKIGPLVDQIYRFSEEGMVNARFDYTYHDSNFRIASIK
PIISETPLVDLYRYDEISGKVEHFGKGVIIYDIITTAVTLSKHFDTHGRIKE
VOYEMFRSLMYWMTVQYDSMGRVTKRELKLGYPYANTTKYTYDYDGDGQLQSVAVNDRP
TWRYSYDLNGLNHLNPGNSVRLMPLRYDLRDRITRLGDIPIYKIDDDGFLCQRGSDVF
EYNSKGILTRAYNKANGWNVQRYRVDGLGRRASCCTNLGHLHQYFYADLHNPTRVTHVY
NHSNSEITSLYYDLQGHLFAMESSEGEYYVASDNTGTPLAVFSINGLMTKQLQYTAY
GEIYVDSNPDFQLVIGFHGGLYDPLTKLVHFTQDYDVLAGRWTSPDYTMWKNIGREP
APFNLYMFKSNNNPLSNELDLKNYVTDVKSWLVMFGFQLSNIIPGFPRAKMYFVSPPYE
LTESQACENGQLITGVQQTTERHNOAFMALEGQVISKRLHASIREKAGHWFATSTPII
GKGIMFAVKKGRVTTGISSIATDDSRKIASVLNSAHYLEKMHYSIEGKDTHYFVKIGS
ADSDLVLTAMTSGRKVLDSGVNVTVSQPTLLINGRTRFRFTNIEFOYSTLLINIRYGLT
ADTLDEKARVLDOARQALGSAWAKEQQKARDREGSRVWTDGKQQLLNTGRVQGY
EGYVLPVEQYPELADSSSNIQFLRQNMGMKR"

mat_peptide

<1..>8406
/gene="TEN2"
/product="teneurin-2"
/evidence=experimental

ORIGIN

Alignment Scores:

Pred. No.: 4.9e-61 Length: 8409
Score: 625.00 Matches: 113
Percent Similarity: 90.07% Conservative: 23
Best Local Similarity: 74.83% Mismatches: 15

Query Match: DB:	79.92% 5	Indels: Gaps:	0 0
US-10-029-020-14_COPY_1250_1400 (1-151) x GGA279031 (1-8409)			
QY	1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20
Db	3856	TACATTGGCGGTATCTTCCCATCCAGGAATGTGACTAGCATATTTGGAGCTGAGAAATAAA	3915
QY	21	AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer	40
Db	3916	GAGTTTAAACATAGCAACAATCCTGCTCACAAATACTATCTGGCCGTGGACCCCGTTTCG	3975
QY	41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60
Db	3976	GGCTCCCTGTACGTATCAGACACCAACAGCCGACGGATATACAAAGTCAAATCTCTTACT	4035
QY	61	ValVallLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu	80
Db	4036	GGCAGAAAGACCTGGCTGGTAATTTCTGAAGTGGTAGCGGGACTGGAGAGCAATGCCTG	4095
QY	81	ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn	100
Db	4096	CCCTTTGATGAAGCCAGATGTGGAGATGGAGGAAAGCAGTGGACCAACCTAATGAGT	4155
QY	101	ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile	120
Db	4156	CCTCGAGGAATTGCAGTGGATAAGTATGACTCATGTATTTTGTGTGATGCCACTATGATT	4215
QY	121	ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer	140
Db	4216	CGAAAAGTGGATCAGAATGGAATTATATATCAACTCTGTGGCTCCAATGACCTAACTGCC	4275
QY	141	AlaArgProLeuSerCysAspSerValMetAsp	151
Db	4276	GTCCGACCTCTAAGCTGTGATTCACGATGGAT	4308
RESULT 7	AX250068 8409 bp DNA linear PAT 28-SEP-2001		
AX250068	Sequence 67 from Patent WO0166747.		
LOCUS	AX250068		
DEFINITION	AX250068		
ACCESSION	AX250068.1 GI:15864504		
VERSION	Gallus gallus (chicken)		
KEYWORDS	Gallus gallus		
SOURCE	Gallus gallus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
REFERENCE	1 Vernet,C.A., Fernandes,E., Shimkets,R.A., Herrmann,J.L., Majumder,K., Macdougall,J., Mishra,V., Mezes,P.S. and Rastelli,L. Proteins named fctrx and nucleic acids encoding same Patent: WO 0166747-A 67 13-SEP-2001;		
AUTHORS	Curagen Corporation (US)		
TITLE	Location/Qualifiers		
JOURNAL	1. .8409		
FEATURES	/organism="Gallus gallus"		
source	/mol_type="unassigned DNA"		
	/db_xref="taxon:9031"		
ORIGIN	Alignment Scores: 4.9e-61 Length: 8409		
	Pred. No.: 625.00 Matches: 113		
	Score: 90.07% Conservative: 23		
	Percent Similarity: 74.83% Mismatches: 15		
	Best Local Similarity: 79.92% Indels: 0		
	Query Match: 6 Gaps: 0		
DB:	US-10-029-020-14_COPY_1250_1400 (1-151) x AX250068 (1-8409)		
QY	1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20

```
/clone="TESTI4009028"
/tissue type="testis"
/clone_lib="TESTI4"
/note="cloning vector: pME18SFL3"

ORIGIN
Alignment Scores:
Pred. No.: 7.99e-61 Length: 3394
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AK125869 (1-3394)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 1964 TAGTGGCGGGATATCCCTTCTGGAATGTAACAAGTGCTTAGAACTAAGAAATAA 2023

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPrometSer 40
Db 2024 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 2083

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 2084 GGAGATCTGTACGTTTCTGACACAAACACCCGCAGAAATTTATCGCCCAAGTCACCTTACG 2143

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 2144 GGGGCAAAAGACTTGACTAAAAATGCAGAGTCGTCGAGGGACAGGGAGCAATGCCTT 2203

QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 2204 CCGTTTGACGAGCGGAGATGTGGGATCGAGGAAGCCCGTGGAAGCCACACTCATGAGT 2263

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile 120
Db 2264 CCCAAGGAATGGCAGTTGATAAGAATGGATTAAATCTACTTTGTGATGGAACCATGATT 2323

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2324 AGGAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 2383

QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 2384 GCCAGACCTTTAACTTGTGACACCAGCATG 2413

RESULT 9
AB040888
LOCUS AB040888 5309 bp mRNA linear PRI 22-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1455 protein, partial cds.
ACCESSION AB040888
VERSION AB040888.1 GI:7959170
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Nagase,T., Kikuno,R., Ishikawa,K., Hirosewa,M. and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
JOURNAL DNA Res. 7 (2), 143-150 (2000)
MEDLINE 20277482
PUBMED 10819331
REFERENCE 2 (bases 1 to 5309)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
```

URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)

FEATURES
source

Location/Qualifiers
1..5309

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fh16070"
/clone_lib="pBluescriptII SK plus"
1..5309
/gene="KIAA1455"
<1..>5309
/gene="KIAA1455"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1455 protein"
/protein_id="BAA95979.1"
/db_xref="GI:7959171"
/translation="QNQPYCRGLPDPQDIISQLSQSPQQAksFYDRISFLIGSDST
HVIgespFNKSLASvirgQVLTADGTPLIGNVSFFHYPEYGYTITRQDGMFDLVAN
GGASLTIVFERSpFLTQYHTWIPWNVFYVMDTLVMKKEENDIPSCDLSGFVRPNPII
VSSPLSTFRSSPEDSPIIPETQVLHEETTIPGTDLKLsYLSsRAAGYKsVLKITMTQ
SIIPFNLMKVHLmVAVVGRLFQKFPASPnLAYTFIWDKTDAYNQKVYGLSEAVSVG
YEYsCLDLTLWEKRTAILQGYELDASNMGGWTLDKHHVLDVQNGILYKNGENQFIS
QOPVVSIMGNRRRSISCPSCNGQADGNKLLAPVALACGIDGSlyVGDfNVYRRIF
PSGNVTSVLELRNKDFRHSSNPahRYLADpVTGdlyVSDnTRRIYRPKSLTGAKD
LTkNAEVVAGTGEQCLPFdeARCGDGGKAVEATLMSPKGMADKNGLIYfVdGTMIrK
VDQNGIISLlGSNDLTsARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLdNNVV
LQITENRQVRIAGRPmHCQVPGEYpVGKHAVQTTLsATAIAVSySGVLIYITeDE
KKINrIRQVTTDGEISlVAGIPSECdCKNDANCdCYQSGGYAKDAKLsAPSSLAASP
DGTLYIADLGNIRIRAVsKNKPLlNSMNFYEVASPTDQELYIFDINGTHQYTVSLVTG
DLYNfFSYsNDNDITAVTDsNGNTLRIRRDpNRMPVRVVSFDNQVIlWTIGTNGCLKS
MTAQGLELVLFtYHGNSGLLATKSDETGWtTFDyDSEGLTNVTFpTGvVTNLHGDM
DKAITVDIESSsREEDVSITsNLSSIDSFYTMVQDQLRNSYQIGYDGSRLRIIYASGLD
SHYQTEPHVLAGTANPTVAkRNMtLPGENGQNLVEWRfRKEQAQGVNVfGRKLrVNG
RNLLsVDfDRtTKTEKIYDdHRKELLRiAYDTSGHPTLWLPSSKLMAVNVTYSSTGQI
ASIQRTTSEKVDYDQGRiVSRVFADGKTWSYtYLEKsMVLlLLHSQRQYiFEYDMWD
RLsAITMPSVARHTMQTIRSiGYRNIYNPpESNAIITDYNEEGLLlQTAFLTsRR
VLFKYRQTRLSEILYDSTRVSfTYDeTAGVLKTVNLQSDGFICTIRYRQIGPLIDRQ
IFRESEDGMVNARFDYSdNSFRVTSMQGVINETPLpIDLYQfDDISGKVEQFGKFGV
IYyDINQIISTAVMTYTKHfDAHGRIKEIQYEiFRSLMWITiQYDNMGRTVKREIKI
GFANTTKAYEYDVdGQLQTVYLNEKIMWRYNdLNGNLHLLNPNSARLTPLRYDL
RDRITRLGDVQVRLDdGfLRQGTETeIfeYSSKGLLTrVYsKSGWTViYRYDGLGRR
VSSKTSLGQHLQfFYADLTYPTRIthVYNHSSSEITsLYDlQHLfAMEISSGDEFY
IASDNTGTPLAVfSSNGLMlKQIOYtAYGEiYFDsNIDfQLViGFHGGLYDPLTKLIH
FGERDYDILAGRWtTPDIeIWKRIGKDPAPFNfMNNPNPASKIHdVKDYTDVNSW
LVTFEGHLHNAIPGFpVPKFDLTepSYELVKsQQWDDIPPIFGVQQVArQAKAFSL
GKMAEVQVSRRRAGAsWLWFATVKSLIGKVMlAVSQRVQTNVLNIANEDCIKVA
AVLNNAFYLENLHfTIEGKDTHYfIKTTPESDLGTLRlLTSGR"

ORIGIN

Alignment Scores:

Pred. No.: 1.37e-60 Length: 5309
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AB040888 (1-5309)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 1161 TAGTGGCGGGATATCCCTTCTGGAATGTAACAAGTGCTTAGAACTAAGAAATAA 1220

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPrometSer 40

Db 1221 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 1280

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 1281 GGAGATCTGTACGTTTCTGACACAAACACCCGCAGAAATTTATCGCCCAAGTCACCTTACG 1340

```
QY      61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu  80
      1341 GGGGCAAAAGACTTGACTAAAAATGCAGAAAGTCGCGAGGGACAGGGGAGCAATGCCTT 1400
QY      81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
      1401 CCGTTTGACGAGGCGGAGATGTGGGGATGGAGGAAGGCCGTGGAAAGCCACACTCATGAGT 1460
QY     101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      1461 CCCAAAGGAATGGCAGTTGATAAGAATGGATTAAATCTACTTCTGGGCTCTAACGATTTGACTTCA 1520
QY     121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      1521 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 1580
QY     141 AlaArgProLeuSerCysAspSerValMet 150
      1581 GCCAGACCTTTAACTTGTGACACCAGCATG 1610

RESULT 10
AX662357
LOCUS      AX662357      8473 bp      DNA      linear      PAT 22-MAR-2003
DEFINITION Sequence 39 from Patent WO02062999.
ACCESSION  AX662357
VERSION     AX662357.1  GI:29163218
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
JOURNAL     Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
            Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
            Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
            Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
            Millet, I. and Macdougall, J.R.
            Proteins and nucleic acids encoding same
            Patent: WO 02062999-A 39 15-AUG-2002;
FEATURES    Curagen Corporation (US)
            Location/Qualifiers
            source
            1..8473
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      2.42e-60      Length:      8473
Pred. No.:      619.00      Matches:      111
Score:      90.67%      Conservative:      25
Percent Similarity:      74.00%      Mismatches:      14
Best Local Similarity:      79.16%      Indels:      0
Query Match:      6      Gaps:      0
DB:

US-10-029-020-14_COPY_1250_1400 (1-151) x AX662357 (1-8473)

QY      1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys  20
      3600 TACGTGGCGGGATATCCCTTCTGGAATGTAAACAAGTCTTAGAACTAAGAAATAAA 3659
QY     21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer  40
      3660 GATTTAGACATAGCAGCAACCCAGCTCATAGATACTACTTGTGCAACGGATCCAGTCACG 3719
QY     41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal  60
      3720 GGAGATCTGTACGTTTCTGACACAAACACCCGCAGAAATTATCGCCCAAGTCACCTTACG 3779
QY     61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu  80
      3780 GGGCAAAAGACTTGACTAAAAATGCAGAAAGTCGTGCGAGGGACAGGGGAGCAATGCCTT 3839
```

```
QY      81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
      3840 CCGTTTGACGAGGCGAGATGTGGGGATGGAGGAAGGCCGTGGAAAGCCACACTCATGAGT 3899
QY     101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      3900 CCCAAAGGAATGGCAGTTGATAAGAATGGATTAAATCTACTTCTGGTGGAAACCATGATT 3959
QY     121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      3960 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 4019
QY     141 AlaArgProLeuSerCysAspSerValMet 150
      4020 GCCAGACCTTTAACTTGTGACACCAGCATG 4049

RESULT 11
AX662359
LOCUS      AX662359      8487 bp      DNA      linear      PAT 22-MAR-2003
DEFINITION Sequence 41 from Patent WO02062999.
ACCESSION  AX662359
VERSION     AX662359.1  GI:29163219
KEYWORDS    Homo sapiens (human)
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,
JOURNAL     Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,
            Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,
            Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,
            Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,
            Millet, I. and Macdougall, J.R.
            Proteins and nucleic acids encoding same
            Patent: WO 02062999-A 41 15-AUG-2002;
FEATURES    Curagen Corporation (US)
            Location/Qualifiers
            source
            1..8487
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Alignment Scores:      2.42e-60      Length:      8487
Pred. No.:      619.00      Matches:      111
Score:      90.67%      Conservative:      25
Percent Similarity:      74.00%      Mismatches:      14
Best Local Similarity:      79.16%      Indels:      0
Query Match:      6      Gaps:      0
DB:

US-10-029-020-14_COPY_1250_1400 (1-151) x AX662359 (1-8487)

QY      1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys  20
      3596 TACGTGGCGGGATATCCCTTCTGGAATGTAAACAAGTCTTAGAACTAAGAAATAAA 3655
QY     21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer  40
      3656 GATTTAGACATAGCAGCAACCCAGCTCATAGATACTACTTGTGCAACGGATCCAGTCACG 3715
QY     41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal  60
      3716 GGAGATCTGTACGTTTCTGACACAAACACCCGCAGAAATTATCGCCCAAGTCACCTTACG 3775
QY     61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu  80
      3776 GGGCAAAAGACTTGACTAAAAATGCAGAAAGTCGTGCGAGGGACAGGGGAGCAATGCCTT 3835
QY     81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
      3835
```


Db 3836 CCGTTTGACGAGCGGAGATGTGGGGATGGAGGGAAGCCGTGGAAGCCACACTCATGAGT 3895

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120

Db 3896 CCCAAAGGAATGGCAGTTGATAAGAAATGGATTAAATCTACTTTGTTGATGGAACCATGATT 3955

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140

Db 3956 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 4015

QY 141 AlaArgProLeuSerCysAspSerValMet 150

Db 4016 GCCAGACCTTTAACTTTGTGACACCAGCATG 4045

RESULT 12

AX662355

LOCUS

AX662355

DEFINITION

Sequence 37 from Patent WO02062999.

ACCESSION

AX662355

VERSION

AX662355.1 GI:29163217

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

REFERENCE

1

AUTHORS

Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X., Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K., Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A., Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F., Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G., Millet, I. and Macdougall, J.R.

TITLE

Proteins and nucleic acids encoding same

JOURNAL

Patent: WO 02062999-A 37 15-AUG-2002;

FEATURES

source

1. 8645

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2.48e-60 Length: 8645

Score: 619.00 Matches: 111

Percent Similarity: 90.67% Conservative: 25

Best Local Similarity: 74.00% Mismatches: 14

Query Match: 79.16% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AX662355 (1-8645)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 3772 TACGTGCGGCGGATATTCCCTTCTGGAAATGTAAACAAGTGTCTTAGAACTAAGAAATAAA 3831

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40

Db 3832 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 3891

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 3892 GGAGATCTGTACGTTTCTGACACAAACACCCGCGAGAATTATCGCCCAAGTCATTACG 3951

QY 61 ValValLysAspLeuValLysAsnSerGluValAlaIleGlyThrGlyAspGlnCysLeu 80

Db 3952 GGGGCAAAAGACTTGACTAAAAATGCAGAGTCGTCGAGGGACAGGGGAGCAATGCCTT 4011

QY 81 PropheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100

Db 4012 CCGTTTGACGAGCGGAGATGTGGGATGGAGGGAAGCCGCTGGAAGCCACATCATGAGT 4071

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120

Db 4072 CCCAAAGGAATGGCAGTTGATAAGAAATGGATTAAATCTACTTTGTTGATGGAACCATGATT 4131

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140

Db 4132 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 4191

QY 141 AlaArgProLeuSerCysAspSerValMet 150

Db 4192 GCCAGACCTTTAACTTTGTGACACCAGCATG 4221

RESULT 13

AX662353

LOCUS

AX662353

DEFINITION

Sequence 35 from Patent WO02062999.

ACCESSION

AX662353

VERSION

AX662353.1 GI:29163216

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

REFERENCE

1

AUTHORS

Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X., Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K., Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A., Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F., Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G., Millet, I. and Macdougall, J.R.

TITLE

Proteins and nucleic acids encoding same

JOURNAL

Patent: WO 02062999-A 35 15-AUG-2002;

FEATURES

source

1. 8675

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2.49e-60 Length: 8675

Score: 619.00 Matches: 111

Percent Similarity: 90.67% Conservative: 25

Best Local Similarity: 74.00% Mismatches: 14

Query Match: 79.16% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x AX662353 (1-8675)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 3784 TACGTGCGGCGGATATTCCCTTCTGGAATGTAAACAAGTGTCTTAGAACTAAGAAATAAA 3843

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40

Db 3844 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 3903

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 3904 GGAGATCTGTACGTTTCTGACACAAACACCCGCGAGAATTATCGCCCAAGTCATTACG 3963

QY 61 ValValLysAspLeuValLysAsnSerGluValAlaIleGlyThrGlyAspGlnCysLeu 80

Db 3964 GGGGCAAAAGACTTGACTAAAAATGCAGAGTCGTCGAGGGACAGGGGAGCAATGCCTT 4023

QY 81 PropheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100

Db 4024 CCGTTTGACGAGCGGAGATGTGGGATGGAGGAAGCCGCTGGAAGCCACATCATGAGT 4083

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120

Db 4084 CCCAAAGGAATGGCAGTTGATAAGAAATGGATTAAATCTACTTTGTTGATGGAACCATGATT 4143

```
QY      121  ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      4144  AGGAAAGTTGACCAAAATGGAATCATATCAACTTCTGGGCTCTAACGATTGACTTCA 4203

QY      141  AlaArgProLeuSerCysAspSerValMet 150
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      4204  GCCAGACCTTTAACTTGTGACACCAGCATG 4233

RESULT 14
AK122513      5804 bp      mRNA      linear      ROD 15-MAR-2003
LOCUS      Mus musculus mRNA for mKIAA1455 protein.
ACCESSION      AK122513
VERSION      AK122513.1 GI:28972757
KEYWORDS      FLI_CDNA.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS      1  Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S.,
      Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
      Prediction of the coding sequences of mouse homologues of KIAA
      gene: II. The complete nucleotide sequences of 400 mouse
      KIAA-homologous cDNAs identified by screening of terminal sequences
      of cDNA clones randomly sampled from size-fractionated libraries
      DNA Res. 10, 35-48 (2003)
      2  (bases 1 to 5804)
      Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
      Direct Submission
      Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research
      Institute, Laboratory for Genome Informatics; 2-6-7
      Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
      (E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
      The CREATE program supported by Japan science and technology
      corporation; cDNA full insert sequencing: Kazusa DNA Research
      Institute; cDNA library construction, clone selection and 5'- &
      3'-end one pass sequencing.
      Location/Qualifiers
      1. .5804
      /organism="Mus musculus"
      /mol_type="mRNA"
      /db_xref="taxon:10090"
      /clone="mbg00055"
      /tissue_type="brain"
      /dev_stage="adult"
      /note="vector:modified pBC SK+"
      1. .5804
      /gene="mKIAA1455"
      <1. .5488
      /gene="mKIAA1455"
      /note="CDS is predicted by in silico analysis. Start codon
      is not identified."
      /codon_start=2
      /evidence=not experimental
      /product="mKIAA1455 protein"
      /protein_id="BAC65795.1"
      /db_xref="GI:28972758"
      /translation="LIGNVVSFLHYSEYGYTITRQDGMFDLVANGASLTLVFERSPF
      LTQYHTVWPWNVFYVMDTLVMKKEENDIPSCDLGFRPSPPIVSSPLSTFFRSSPE
      DSPIIPETQVLIHEETTIPGTDLKL SVLSSRAAGYKSVLKITMTQAVIPFNLMKVHLMV
      AVVGRLFQWFPASPNLAYTFIWDKTDAYNQKYGVLSEAVSVSGYEYESCLDLTLWEK
      RTAVLQGYELDASNMGGWTLDKHHVLDVQNGILYKNGENQFISQOPPVVSSIMNGR
      RRSISPCSCNGQADGNKLLAPVALACGIDGSLYVGDFNYVRRIFPSGNVTSVLELRNK
      DFRHSSNPAHYLYLATDPVTGDLVSDTWTTRRIYRPSKLTGAKDLTKNAEVVAGTGEQ
      CLPFDEARCGGKAVEATLMSPKGMAIDKNGLTYFVDGTMIRKVDQNGIISTLLGSN
      DLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDNNVVLQITENRQVRIAAG
      RPMHCQVPGEYVPGKHAVQTTLESATAIAYSGVLTYITETDEKKNIRIQVTTDGE
      ISLVAGIPSECDCKNDANCDCYQSGDGYAKDAKLNAPSSLAASPDGTLIADLGNIRI
      RAVSKNKLPLINSMNPFYEVASPTDQELYIFDINGTHQYTVSLVTGDLNLYNFSYSNNDV
      TAVTDSNGNTLRIIRDPNRMPVRVSPDNQVITWTGTGNCCLKSMTAQGLELVLTFTYH
      GNSGLLATKSDETGWTTFDYSDEGRLTNVTFPTGNDMDKAITVDIESSSRE
      EDVSIITSNLSSIDSFYTVMVQDQLRNSYQIGYDGLRIFYASGLDLSHYQTEHVLAGTA
```

```
NPTVAKRNMTLPGENGQNLVEWRFRKEQAQGVNVFGRKLRVNGRNLISVDFDRTTKT
EKIYDDHRKFLLRITAYDTSGHPTLWLPSSKLMANVNTSYSTGQIASIQRGTTSEKVDY
DSQGRIVSRVFADGKTWSYTYLEKSMVLLHSQRYIIFYEDMDRLSALTMPSPVARHT
MQTIRSIGYVRNIYNPPESNAIITDYNEEGLLQTAFLGTSRRVLFKYRRQTRLSEI
LYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIGPLIDRQIFRPSDEGVMVNARF
DYSYDNSFRVTSMQGVINETPLPIDLYQDDISGKVEQFGKFGVIYDINQIISTAVM
TYTKHFDAGRKEIKIYEIFRSLMYWITIYDNMRGRVTKREIKIGPEFANTTKYAYEYD
VDGQLQTVYLNKIMWRNYDNLNGLHLLNPPSSSARLTPLRYDLRDRITRLGDVQYRL
DEDGFLRQRTFIEYSSKGLLTRYVSKGSGWTVIYRYDGLGRRVSSKTSLGQHLQFF
YADLTYPTRITHVYNHSSSEITSLYYDLQGLHFAFMEISSGDEFYIASDNTGTPLAVFS
SNGMLKQIQIQTAYGEIYFDSNVDFQVIGFHGGLYDPLTKLIHFGERYDIDLAGRWT
TPDIEIWKRIKGPAPFNLYMFRNNPNPASKIHDVKDYITDVSNLVTFGFHLHNAIPG
FPVPKFDLTPESEYELVKSQWEDVPPIFGVQOQVARQAKAFSLGKMAEVQVSRRKAG
AEQSWLWFATVKSLLIGKVMIAVSGQRVQTNVNIANEDCIKVAAVLNNAFVLENLHF
TIEGKDTHYFIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVNGRTRFPADVE
MOFGALALHVRYGWMTLDEEKARILEQARQALARAWAREQQORVRDGEEGARLWTEGK
ROLLSAGKVQGYDGYVLSVEQYPELADSANNIQFLRQSEIGKR"
```

ORIGIN

```
Alignment Scores:      1.99e-60      Length:      5804
Pred. No.:      618.00      Matches:      110
Score:      90.67%      Conservative:      26
Percent Similarity:      73.33%      Mismatches:      14
Best Local Similarity:      79.03%      Indels:      0
Query Match:      10      Gaps:      0
DB:

US-10-029-020-14_COPY_1250_1400 (1-151) x AK122513 (1-5804)

QY      1  TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      944  TACGTCGGCGGATATTCCCGICTCGGAATGTGACAAAGTGTTAGAACTAAGAAATAA 1003

QY      21  AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1004  GATTTTAGACATAGTAGCAACCCAGCTCACAGATACCTGGCTACGGACCCAGTCACC 1063

QY      41  GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1064  GGAGATTTCGTACGTCTCTGTATACTATAACACCCGCAGAAATCTATCGGCCGAAATCACTCACG 1123

QY      61  ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1124  GGAGCCAAAGACCTGACTATAAAACGCTGAAGTGGTGGCAGGACCGGGGACAGTGCCTT 1183

QY      81  PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1184  CCTTTGACGAGGCCAGGTGTGGGGATGGAGGCAAGGCTGTGGAAGCAACGCTCATGAGT 1243

QY      101  ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1244  CCCAAAGGAATGGCAATCGATAAGAACGGGACTGATCTACTTTGTTGATGGAACCATGATC 1303

QY      121  ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1304  AGAAAGGTTTGATCAAAATGGAATCATATCAACTCTCTGGGCTCCAACGACCTCACGTCA 1363

QY      141  AlaArgProLeuSerCysAspSerValMet 150
      |||:::|||||:::|||||:::|||||:::|||||:::|||||
Db      1364  GCTCGACCTTTAACCTGTGATAGCATG 1393
```

RESULT 15

```
AB025412      8964 bp      mRNA      linear      ROD 08-MAY-1999
LOCUS      Mus musculus mRNA for Ten-m3, complete cds.
DEFINITION      AB025412
ACCESSION      AB025412
VERSION      AB025412.1 GI:4760779
KEYWORDS      Ten-m3.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1  (sites)
```

AUTHORS

Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.

TITLE

Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues

JOURNAL

J. Cell Biol. (1999) In press

REFERENCE

2 (bases 1 to 8964)

AUTHORS

Oohashi,T.

TITLE

Direct Submission

JOURNAL

Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho, Okayama, Okayama 700-8252, Japan

REFERENCE

(E-mail:oohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128, Fax:+81-86-222-7768)

FEATURES

source

1. .8964

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Balb/c"

/db_xref="taxon:10090"

/tissue_type="brain"

/dev_stage="adult"

1. .8964

/gene="ten-m3"

50. .8197

/gene="ten-m3"

/codon_start=1

/product="Ten-m3"

/protein_id="BAA77398.1"

/db_xref="GI:4760780"

translation="MDVKRRPYCSLTKSRREKERRYNTSSADNEECRVPTQKSYSSS

ETLKAFDHDYSLLYGNRVKDLVHREADEYTRQGNFTLRQLGVCESATRRGVAFCAE

MGLPHRGYSISAGSDADTENEAVMSPEHAMRLWGRGVKSGRSSCLSSRSNSALTLDIT

EHENRSDSESEQPSNPNPGQTLQPLPPSHKHQHPAQHPHSITSLNRNSLTNRNQSPAP

PAALPAELQTTPTESVQLQDSWVLGSNVPLESRHFLFKTGTGTTPLFSTATPGYTMASG

SVYSPPTRLPRLNTLSRAFKFKKSSKYCSWRCTALCAVGSVLLAILLSYFIAMHLF

GLNWHLQQTENDTTFENGKVNSTVPTNTVSLPSGDNGLKGGFTHENNTIDSGELDIGR

RAIQEVPPGIFWRSQLFIDQPQLKFNISLQKDALIGVYGRKGLPPSHDQYDFVELLD

GSRLIAREQNLVESERAGQARSVSLHEAGFIQYLDSGIWHLAFYNDGKNBQVSNFN

TIVIESVECPNRCHNGECVSGTCHFPGLGPCDSRAACPVLCSGNGQVSKGRCLC

FSGWKGTECDVPTTQCIDPQCGGRCICIMGSCACNSGYKGENCEEADCLDPGCSNHGV

CIHGECNPNPGWGSNIEILKTMCADQCDSGHGTVLQESGCTCDPNWTGPDCSNEICS

VDCGSHGVCMGSCRCHEGWTGPACNQRACHPRCAEBHTCKDGKCECSQGWNGEHCTI

AHYLDKIVKEGPGCLNSNGRCTLDQNGWHCVCPQCMRGAGCDVAMETLCTDSKDNEG

DGLIDCMDPDCCLOSSCQNPYCRGLDPQDIISQSLQTPSQQAASKFYDRISFLIGS

DSTHVLPGESPENKSLASVIRQVLTADGTPLIGVNVSVFLHYSEGYTITRQDGMFDL

VANGGASLTLVFERSPFLTQYHTVWIPWNVFIYMDTLVMKKEENDIPSCDLSGFVRPS

PIIVSSPLSTFFRSSPEDSPIIPETQVLHEETITPGTDLKL SylSSRAAGYKSVLKIT

MTQAVIPENLMKVHLMAVVGRFLQKWFPPSPNLAYTIFWDKTDAYNOKVYGLSEAVV

SVGYEYESCLDLTLWEKRTAVLQGYELDASNMGWTLDKHHVLDVQNGILYKNGENQ

FISQQPPVSSIMNGRRRSISCPSCNGQADGNKLLAPALACGIDGSLYVGDFENYVR

RIFPSGNVTSVLELRNKFDRHSSNPAHYLYLATDPVTGDLYVSDTNTRIYRPKSLTG

AKDLTKNAEIVVAGTGEQCLPFDEARCGDGKGAVEATLMSPKGMAIDKNGILYFVDGTM

IRKVDQNGIISLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDN

NVVLQITENQVRIIAAGRPMHCQVGEYVPGKHAVQTILESATAIAVSYSGVLYITE

TDEKKINRIRQVTTDGEISLVAGIPSECDCKNDANCDCYQSGDGYAKDAKLNA PSSLA

ASPDGTLXTADLGNIRIRAVSKNKPILNSMNFYEVASPTDQELYIFDINGTHQYTVSL

VTGDYLYNFSYNDNDVTAVTDNSGNTLRIRRDPNRMPVRVSPDNQVIWLITIGTNGC

LKSMTAQGLLELVLFYHNSGLLATKSDETGTTTFDDYDSEGRLTNVTFTGVTNLH

GMDKAITVDIESSSREEDVSI TSNLSSIDSFTVMQDLRNSYQIGYDGLSRLIFYAS

GLDSHYQTEPHVLAGTANPTVAKRNMILPGENGONLVWEFRKEQAQKVNVPFRKLR

VNGRNLLSVDFDRTTKTEKIYDDHRKFLRIAYDTSGHPTLWLPSSKMAVNVVTSST

GQIASIQRTTSEKVDYDSQGRIVSRVFADGKTWSYTYLEKSMVLLLSQRQYIFEYD

MWDRLSAITMPSVARHTMQTIRSIGYVRNINYPPEASNASII TDVNEEGLLLQTAFLGT

SRRLVFKYRQTRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYQIGPLI

FRQIFRSEDGMVNARFDYSYDNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGK

DGVIYYDINQIISTAVMTYTKHFDAGRIKEIQEYIFRSLMWITIQYDNMGRVTKRE

IKIGPEFANTKYAYEYDVVDGQLQTVVLNKEIMWRYNVLDNGLNLHLLNPFSSARLTPLR

YDLRDRITRLGDVQYRLDEDEGFLRQGTETFEYSSKGLLTRVYSKSGSGWTVIYRYDGL

GRRVSSKTSLGQHLQFFYADLTYPTRITHVYNHSSSEITSLYYDLQGHLFAMEISSGD

EFYIADNTGTPLAVFSSNGLMKOTQYTAIGEIIYFDSNVDFQLVIGFHGGLYDPLTK

LIHFGERDYDILAGRWTTPDIEIKWRIGKDPAPPNLYMFRNPNPASKIHDKVQYITDV

NSWLVTFGFHLHNAIPGFPVPKFDLTPEPSYELVKSQQWEDVPPIFGVQQQVARQAKAF

LSLGKMAEVQVSRRKAGAEQSWLWFATVKSLIGKGVMLAVSQGRVQTNVLNIANEDCI

ORIGIN

Alignment Scores:

Pred. No.:

3.37e-60

Length:

8964

Score:

618.00

Matches:

110

Percent Similarity:

90.67%

Conservative:

26

Best Local Similarity:

73.33%

Mismatches:

14

Query Match:

79.03%

Indels:

0

DB:

10

Gaps:

0

US-10-029-020-14_COPY_1250_1400 (1-151) x AB025412 (1-8964)

QY

1

TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys

20

Db

3653

TACGTCCGGCGGATATTCCCGTCTGTGGAAATGTGACAAAGTGTTTTAGAACTAAGAAATAAA

3712

QY

21

AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer

40

Db

3713

GATTTTAGACATAGTAGCAACCCAGCTCACAGATACTACCTGGCTACGGACCCAGTCACC

3772

QY

41

GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal

60

Db

3773

GGAGATTGTACGTCTCTGATACTATAACACCCGCAGAAATCTATCGGCCGAAATCACTCAGC

3832

QY

61

ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu

80

Db

3833

GGAGCCAAAGACCTGACTAAAAACGCTGAAGTGGTGAGGACCGGGGAACAGTGCCTT

3892

QY

81

ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn

100

Db

3893

CCCTTTGACGAGGCCAGGTGTGGGATGGAGGCAAGGCTGTGGAAGCAACGCTCATGAGT

3952

QY

101

ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle

120

Db

3953

CCCAAAGGAATGGCAATCGATAAGAACCGGACTGATCTACTTTGTTGATGGAACCATGATC

4012

QY

121

ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer

140

Db

4013

AGAAAGGTTGATCAAAATGGAATCATATCAACTCTCCTGGGCTCCAACGACCTCACGTCA

4072

QY

141

AlaArgProLeuSerCysAspSerValMet

150

Db

4073

GCTCGACCTTTAACCTGTGATACTAGCATG

4102

Search completed: August 14, 2004, 11:52:42
Job time : 2702.71 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 4616.26 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750
Perfect score: 1384
Sequence: 1 KINRIRQVTTSGEISLVAGA.....SSKDDVTITNLSASGAFYT 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1384	100.0	8354	6	AX556500	AX556500 Sequence
2	1384	100.0	8438	6	AX675551	AX675551 Sequence
3	1384	100.0	8993	9	HSM806812	BX640737 Homo sapi
4	1380	99.7	8645	6	AX600210	AX600210 Sequence
5	1370	99.0	8585	10	AB025413	AB025413 Mus muscu
6	1361	98.3	9722	10	AF059485	AF059485 Mus muscu
7	1203.5	87.0	8624	9	AB037723	AB037723 Homo sapi
8	1130	81.6	3048	9	AK127705	AK127705 Homo sapi
9	1112	80.3	112872	2	AP000716	AP000716 Homo sapi
10	1112	80.3	165812	2	AP001141	AP001141 Homo sapi
11	1112	80.3	170156	2	AP002412	AP002412 Homo sapi
12	1112	80.3	186084	9	AP002768	AP002768 Homo sapi
13	1103	79.7	9264	5	AB026980	AB026980 Danio rer
14	1098	79.3	226256	10	AC073599	AC073599 Mus muscu
15	1091	78.8	245722	2	AC120288	AC120288 Rattus no
16	1091	78.8	257693	2	AC103323	AC103323 Rattus no
17	943.5	68.2	5804	10	AK122513	AK122513 Mus muscu
18	943.5	68.2	7816	10	AF195418	AF195418 Mus muscu
19	943.5	68.2	8964	10	AB025412	AB025412 Mus muscu
20	939.5	67.9	5309	9	AB040888	AB040888 Homo sapi
21	935.5	67.6	8473	6	AX662357	AX662357 Sequence
22	935.5	67.6	8487	6	AX662359	AX662359 Sequence
23	935.5	67.6	8645	6	AX662355	AX662355 Sequence
24	935.5	67.6	8675	6	AX662353	AX662353 Sequence
25	935	67.6	5583	10	AK122490	AK122490 Mus muscu
26	917	66.3	8118	5	GGA238613	AJ238613 Gallus ga
27	914	66.0	8373	10	AB025410	AB025410 Mus muscu
28	898	64.9	8297	9	AF100772	AF100772 Homo sapi
29	889	64.2	193544	5	AL807754	AL807754 Zebrafish
30	884.5	63.9	8816	5	AB026979	AB026979 Danio rer
31	865.5	62.5	8689	6	AX250067	AX250067 Sequence
32	865.5	62.5	8689	10	AF086607	AF086607 Rattus no
33	864.5	62.5	8797	6	AX250063	AX250063 Sequence
34	864.5	62.5	8797	6	AX250066	AX250066 Sequence
35	864.5	62.5	8797	10	AB025411	AB025411 Mus muscu
36	863.5	62.4	6560	6	AX250065	AX250065 Sequence
37	863.5	62.4	7713	9	HSM808325	BX648178 Homo sapi
38	863.5	62.4	7781	9	AB032953	AB032953 Homo sapi
39	863.5	62.4	9729	6	AX250013	AX250013 Sequence
40	863.5	62.4	9826	6	AX250008	AX250008 Sequence
41	862.5	62.3	8409	5	GGA279031	AJ279031 Gallus ga
42	862.5	62.3	8409	6	AX250068	AX250068 Sequence
43	861.5	62.2	8575	6	AX921803	AX921803 Sequence
44	794	57.4	229533	2	AC106382	AC106382 Rattus no
45	794	57.4	321049	2	AC137221	AC137221 Rattus no

ALIGNMENTS

RESULT 1

AX556500
LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO2057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,
Smithson,G., Li,L. and Ji,W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
FEATURES
source
1. 8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.92e-128 Length: 8354
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AX556500 (1-8354)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4502 AAGATCAACCGCATCAGCAGGTCACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 4561
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4562 CCCAGTGGCTGTGACTGTAAATAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGT 4621
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4622 TATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTGGCTGTGTGTGCTGATGGGGAG 4681
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4682 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTATCCGGAAGAACACAGCCTTTC 4741
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4742 CTCACACCCAGAACATGTATGAGCTGTCTTCACCAATTGACCAGGAGCTCTATCTGTTT 4801
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4802 GATACCCCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAC 4861
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4862 TTCACCTACACTGGGGACGGCGACATCACACTCATCACAGACAACAATGGCAACATGGTA 4921
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4922 AATGTCGCCCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGTG 4981
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4982 TACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCACACACAAGGACACGAG 5041
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200

Db 5042 TTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAAAC 5101
QY 201 GlyTirPThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5102 GGATGGACAACATTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCCT 5161
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5162 ACTGGCCAGGTGACGAGTTTCCGAAGTGATACACAGACAGTTTCAGTGCATGTCAGGTAGAG 5221
QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5222 ACCTCCAGCAAGGATGATGTACCATAAACCAACCACTGTCTGCCTCAGGGCGCTTCTAC 5281
QY 261 Thr 261
Db 5282 ACA 5284
RESULT 2
AX675551 AX675551 8438 bp DNA linear PAT 27-MAR-2003
LOCUS Sequence 1 from Patent WO2055704.
DEFINITION AX675551
ACCESSION AX675551
VERSION AX675551.1 GI:293333552
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Padigar,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,
Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
Stone,D.J.
TITLE Proteins, polynucleotides encoding them and methods of using the
same
JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
FEATURES
source
1. 8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 2.96e-128 Length: 8438
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AX675551 (1-8438)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4508 AAGATCAACCGCATCAGCAGGTCACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 4567
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4568 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGT 4627
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4628 TATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTGGCTGTGTGTGCTGATGGGAG 4687
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4688 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACACAGCCTTTC 4747

Db 538 CTCTACGTGGCCGACCTTTGGGAACATCCGAATTTCGGTTTATCCGGAAGAACAAAGCCTTTTC 597

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

Db 598 CTCACACCCAGAACATGTATGAGCTGTCTTCACCAATTGACCAGGAGCTCTATCTGTTT 657

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120

Db 658 GATACACCCGCAAGCACCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAAAC 717

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuLeuThrAspAsnAsnGlyAsnMetVal 140

Db 718 TTCACCTACACTGGGACGGCGACATCACACTCATCACAGACAACAATGGCAACATGGTA 777

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160

Db 778 AATGTCGCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCCAGATGGCCAGGTG 837

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180

Db 838 TACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCACACAAAGGACACGAG 897

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200

Db 898 TTGGCCATGATGACATACCATGGCAATTCCGGCCTTTCGGCAACCAAAAGCAATGAAAAC 957

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220

Db 958 GGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCCTGACAAATGTGACCTTCCCT 1017

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240

Db 1018 ACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGCATGTCCAGGTAGAG 1077

QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260

Db 1078 ACCTCCAGCAAGGATGATGTACACATAACCAACCACTGTCTGCCTCAGGCGCCTTCTAC 1137

QY 261 Thr 261

Db 1138 ACA 1140

RESULT 4

AX600210

LOCUS

DEFINITION Sequence 22 from Patent WO02072830.

ACCESSION AX600210

VERSION AX600210.1 GI:28400252

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S., Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G., Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K., Khare,R. and Walia,N.K.

TITLE Proteins associated with cell growth, differentiation, and death

JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;

FEATURES

Location/Qualifiers

source

1 .8645

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="Incyte ID No: 7488573CB1"

ORIGIN

Alignment Scores:

Pred. No.: 7.69e-128 Length: 8645

Score: 1380.00 Matches: 260

Percent Similarity: 99.62% Conservative: 0

Best Local Similarity: 99.62% Mismatches: 1

Query Match: 99.71% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AX600210 (1-8645)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20

Db 4551 AAGATCAACCGCATCAGGCAGGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 4610

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40

Db 4611 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTTTCTGGAGACGATGGT 4670

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60

Db 4671 TATGCCAAGGATGCCAAAGTTAAATACCCCATCTCTCTGGCTGTGTGTGTGGGGAG 4730

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80

Db 4731 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTC 4790

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

Db 4791 CTCAACACCCAGAACATGTATGAGCTGTCTTACCAATTGACCAGGAGCTCTATCTGTTT 4850

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120

Db 4851 GATACCAACCGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 4910

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140

Db 4911 TTCACCTACACTGGGGACGGCGACATCACACTCATCACAGACAACAATGGCAACATGGTA 4970

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160

Db 4971 AATGTCGCGCGAGACTCTACTGGGATGCCCTCTGGTGGTGGTCCCAGATGGCCAGGTG 5030

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180

Db 5031 TACTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGACCACACAGGACACGAG 5090

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200

Db 5091 TTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAAC 5150

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220

Db 5151 GGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCCTGACAAATGTGACCTTCCCT 5210

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240

Db 5211 ACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGCATGTCCAGGTAGAG 5270

QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260

Db 5271 ACCTCCAGCAAGGATGATGTACCATTAACCAACCTGTCTGCCTCAGGCGCCTTCTAC 5330

QY 261 Thr 261

Db 5331 ACA 5333

RESULT 5

AB025413

LOCUS

DEFINITION Mus musculus mRNA for Ten-m4, complete cds.

ACCESSION AB025413

VERSION AB025413.1 GI:4760781

KEYWORDS Ten-m4.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

AB025413 8585 bp mRNA linear ROD 08-MAY-1999

1 (sites)
Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8585)
Oohashi,T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama 700-8252, Japan
(E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)

Location/Qualifiers
1. .8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1. .8585
/gene="ten-m4"
188. .8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translation="MDVKERKPYRSLTRRRDAERRYTSSADSEKGPQKSYSSSET
LKAYDQARLAYGSRVKMVQEAEEFCRTGNTFLRELGLGEMTPPHGTLYRTDIGL
PHGYSMGASSDADLEADTVLSPEHPVRLWGRSTRSGRSSCLSRANSNLTLTDTEHE
NTEIDHPSSLQNHPRLRTPPPPLPHAHTPNQHHAAINSLNRGNTPRSNPSPAPTDH
SLSGEPAGSAQEPHTAQDNWLLNSNIPLERNLKGQPFGLTQDNLIEMDILSARSH
DGAYSDHFLFKPGGTSPLFCTTSPGYPLTSTSVYPPRPLPRSTFSRPAFNLKKPS
KYNWKAALSAILISATVILLAYFVAMHLFGLNWHLQPMEGOMQMEYETEDTASSW
PVPTDVSLYPSGGTGLETPDRKGAAEGKPSLFPEDSFIDSGEIDVGRRASQKIPP
GTFWRSQVFIDHPVHLKFNVSLGKAAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQE
ARSLGPOQRSGPVPPSSSHETGFIQYLDSGIWHLAFYNDGKESVSVFLTTAIESVD
NCPSNCYNGDCISGTCFLGFLGPDGCRASCPVLCSGNQYMKGRCLCHSGWKGA
CDVPTNQCIDVACSSHGTCIMGTCICNPYKGESCEEVDCMDPTCSSRGVCVRGEC
SVGWGTCETPRATCLDQCSHGTFLPDTGLNCNDPSWTGHDCSIEICAADCGGHV
CVGTCCREDGWMGAACDQRAQCHPRCAEHGTCRDGCKECPGWNGEHCTIAHYLDRV
KEGCPCLGNGRCITDLNGWHVCVQLGWRTGCTDMSCTGCGDKNDGDLVDCMD
PDCCLQPLCHVNPLCLGSPDLDIIQETQAPVSQQNLNPFYDRIKFLVGRDSTHSIP
ENPFDDGHACVIRGQVMTSDGTPLVGVNISFINNPLFGYTI SRQDGSFDLVTNGGISI
ILRFERAPFITQETLWLPWDRFFVMEITVMRHEENI PSCDLSNFARPNPVVSPSL
TSFASCAEKGPVPEIQALQBEI VIAGCKMRLSYLSSTRTPGYKSVLRISLTHPTPE
NLMKVHLMAVEGRLEFRKFAAAPDLSSYFIWDKTDVYNOKVFGFEAFVSVGYEYES
CPDLILWEKRTAVLQYEIDASKLGGWLDKHHALNIQSGILHKGNGENQFVSQQPPV
IGSIMNGRRRSISPCSCNGLADGNKLLAPVALTCGSDGLYVGDFNYIRRIFFPSGNV
TNILEMRNKDFRHSHPAHXYLATDPMGSAVFLSDTNRKRVFKVTKVTWKDLVKNS
EVVAGTGQCLPFDPTFCGDKGKATEALTNPGRGITVDKFGLIYFVDGTMIRRVDDNG
IIISTLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSILVLDNNVLOISE
NHQVRI VAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKIN
RIRQVTSGEI SLVAGAPSGCDCKNDANCDPFGDDGAKDAKNTLPSLAVCADGEL
YVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFDTSKKHLYTQSLTDGYLY
NFTYTGDDI THITDNNGNMNVNRDSTGMPLWLVPDQVYVWVTMGTSNALSRSVTTQ
GHELAMTYHGNSGLLATKSNENGWTTFFEYDSFGRLTNNVFTPTQGVSSFRSDTSSV
HVQVETSSKODVTITNLASGAFYTLQDQVRNSVYIGADGSLRLLLANGMEVALQT
EPHLLAGTVNPTVGRNVTLPI DNLGLNQLVWRQESPPSLRGVNVNVTYSPGGHIAQIR
LD FDRVTRTEKI YDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVNVTYSPGGHIAQIR
GIMSERMEYDQAGRTSRI PADGKMWSYTYLEKSMVLHLSQRQYIFEFDKNDRLSSV
TMENVARQTTLETIRSVGYRNIYQPPEGNASVIQDFTEDGHLLHTFYLGTGRRVYIKY
GKLSKLAETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCTIRYRQIGPLIDRQIFRFT
EEGMVNARFDYNDNSFRVTSMAVINETPLPIDLYRVDYDSGKTEQFGKFGYIYVDI
NQIIITTAVMTHTKHFDAYGRMKEVQYEI FRSLMYMTVQYDNMGRVVKELKVGVPYAN
TTRYSYEYDADGQLQTVSINDKPLWRYSDNLGNLHLLSPGNSARLTPLRYDLRDRIT
RLGDVQYKMDGELRQRGDVFYEYNSAGLLIKAYNRA SGWSVRYRYDGLGRRVSSKS
SHSHHLOFFYADLTNPTKTVHLYNHSSEITSLYDYLQGHFLAMELSSGDEFYACDN
IGTPLAVFSGTGLMIKQILLYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRD
YDVLAGRWTSPDHELWKRLSSNISVPFHLMYFNKNPNISNSQDI KCFTMDVNSWLI.TF

RESULT 6
 AF059485 LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998
 DEFINITION Mus musculus Doc4 (Doc4) mRNA, complete cds.
 ACCESSION AF059485
 VERSION AF059485.1 GI:3170614
 KEYWORDS .
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 9722)
 AUTHORS Wang, X.-Z., Kuroda, M., Sok, J., Batchvarova, N., Kimmel, R., Chung, P.,
 Zinszner, H. and Ron, D.
 TITLE Identification of novel stress-induced genes downstream of chop
 JOURNAL EMBO J. 17 (13), 3619-3630 (1998)
 MEDLINE 98315054
 PUBMED 9649432
 REFERENCE 2 (bases 1 to 9722)
 AUTHORS Wang, X.-Z. and Ron, D.
 TITLE Direct Submission
 JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University
 Medical Center, 550 First Ave., New York, NY 10016, USA
 FEATURES
 source
 1. .9722
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NIH/Swiss"
 /db_xref="taxon:10090"
 /cell_line="NIH-3T3"
 /note="tunicamycin-treated"
 1. .9722
 /gene="Doc4"
 /note="downstream of CHOP 4"
 583. .9060
 /gene="Doc4"
 /note="similar to Drosophila melanogaster tenm/odz and
 human gamma-herregulin; type II transmembrane protein"
 /codon_start=1
 /product="DOC4"
 /protein_id="AAC31807.1"
 /db_xref="GI:3170615"
 /translation="MDVKERKPYRSLRRRDAERRYTSSSADSEBGPQKSYSSSET
 LKAYDQARLAYGSRVDMVPQEAEEFCRTGTNFTLRELGLGEMTPPHGTLYRTDIGI
 PHCYMGASDADLEADTVLSPHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTEHE
 NTETGAPLHSSASSTPIEQSPSPPPANESQRRLLGNVGAQTPDSDSEEFVFN
 SFLVKSASLGVAAANDHPSSLQNHRLRTPPPPLPHAHTPNQHHAAASINSLNRGNFT
 PRNSPAPTDLHSLGEPAGSAQEPHAQDNWVLSKI PVETRNLGKQPPFLGTWQDN
 LIEMDIFASRRDGYSDGHFFKPGTSPLECTSPGYLTSSTVYSPPPRPLRST
 FSRPAFNLKPSKYCNWKCAALSAILTSATLIVLAFVAMHLFGLNHWLQPMEGQM
 MYEITETASSWPVPTDVSLYPSGGTGLTETPD RKGKGAEGKPSLSLPEDSFIDSGEI
 DVGRRASQKIPPGTFWRSQVFI DHPVHLKFNVSLGAALVGIYGRKGLPPSHTQLDFV
 ELLDGRRLITQEARSLGEPQQRSGRPVPPSSHETGFIQYLDGSIWHLAFYNDGKESV
 VSLTLTIESVDNCPNCGNGDCISGTCHEFLGFLGDCGRASCPLVCSGNGQYMKG
 RCLCHSGWKAEGCDVPTNQCIDVACSSHGTCIMGTCICNPGYKGESCEEVDCMDPTCS
 SRGVCVRGECHCSVGGWGTNCETPRATCLDQCSGHGTFLPDTGLCNCDPSPWTGDCSI
 EICAADCGHGVCGVGTCTCEDGWMGAACDQACHPRAEHGTCRDGKCECTPGWNGE
 HCTIEGCGLNGNGRCTLDLNGWHCVQLGWRGTGCDTSMETGCGDKDNDGDGLV
 CMDPDCCLQPLCHVNPCLGLSPDLDIIQETQAPVSQNLNSFYDRIFKFLVGRDSTHS
 IPGENPFDGHACVIRGOVMTSDGTPLVGNISFINNPLFGYTI SRQDGSFDLVTNG
 ISIIILRERAPFTQTEHTLWLPWDRFFVMEITVMRHEENEI PSRDLSNFARPNPVVSP
 SPLTSFASSCAEKGPVPEIQALQEEI VIAGCKMRLYSLSRTPGYKSVVRISLTHPT
 I PFNLMKVHLMVAVEGR LFRKWFAAAPDL SYFIWDKTDVYNQKVFGLSEAFVSVGVE
 YESCPDLILWEKRTAVLQGYEIDASKLGGWSLDKHHALNIQSGILKHNGENQFVSQ
 PPVIGSIMNGRRRSICPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFPS
 GNVNTNILEMHSHPAKYYLATDPMSGAVFLSDTNSRRRVFKVKSTTVVKDLVKNSEVA
 GTGDQCLPFDTRCGDGKATEATLTNPRGTVDFKGLIYFVDGTMIRRVQNGLIIST
 LLGSDNLT SARPLSCDSVMEISQVRLEWPTDLAINPMDNSLVLDNNVVLQISENHQV
 RIVAGRPMHCQVPQIDQFLLSKVAIHATLESATALAVSHNGVLVIAETDEKKINRIRO
 VTTSGEISLVAGAPSGCDCKNDANCDFSGDGAYAKDAKLTNPSLAVCADGELYVAD
 LGNIRIRFIRKNKPFNTQNNMYELSSPIDQELYLFDTSKHLTYQSLPTGDIYLNFTY
 TCGDGI THITNNNGMNVNVRDSTGMPI.WI.VVDPGVWVYMTGTNSALRSVYTTQGH
 EL

QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
|||||
Db 5818 GGGTGGACAACGTTTATGAGTATGACAGTTTGGTCGCCTGACAAACGTGACCTTTCCA 5877

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
|||||
Db 5878 ACTGGCCAGGTGACAGTTTCCGAAGCGATACAGACAGCTCAGTGCACGTGCAGGTAGAG 5937

QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
|||||
Db 5938 ACCTCAAGCAAGATGACGTACCATACCCACCTGTCTGTTTCGGGTGCCTTCTAC 5997

QY 261 Thr 261
|||
Db 5998 ACC 6000

RESULT 7
AB037723
LOCUS AB037723 8624 bp mRNA linear PRI 10-MAY-2002
DEFINITION Homo sapiens mRNA for KIAA1302 protein, partial cds.
ACCESSION AB037723
VERSION AB037723.2 GI:20521827
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M., and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVI. The complete sequences of 150 new cDNA clones from brain which
code for large proteins in vitro
JOURNAL DNA Res. 7 (1), 65-73 (2000)
MEDLINE 20181126
PUBMED 10718198
REFERENCE 2 (bases 1 to 8624)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
On May 9, 2002 this sequence version replaced gi:7242958.
COMMENT Location/Qualifiers
FEATURES
source
1..8624
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bf00104"
/tissue_type="brain"
/clone_lib="pBCSKPGW"
/note="This sequence was replaced that of fg06864 cDNA as
a representative cDNA sequence for KIAA1302."
1..8624
/gene="KIAA1302"
/gene="KIAA1302"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1302 protein"
/protein_id="BAA92540.2"
/db_xref="GI:20521828"
/translation="STERLDNKPETTRYDSFGRLLTNVFPPTQVSSFRSDTSSVHVQ
VETSSKDDVITITNLASGAFYTLQDVNSYYIGADGSLRLLLLANGMEVALQTEPH
LLAGTVNPTVKRNVLPIDNGLNLVEMQORKEQARGQVTFGRRLRVHNRNLLSLDF
DRVTRTEKIYDDHRKFTLRILYDQGRPSLWSPSSRLNGVNVTVSPGYIAGIQGIM
SERMEYDQAGRTSRIFADGKTWSYTYLEKSMVLLHSQRYIFEFDKNDRLSSTMP
NVARQLETIRSVGYRYIYQPPENASVIQDFTEDGHLHTFYLGTGRRVYKYKGL
SKLAETLYDTTKVSFTYDETAGMLKTNLQNEGFTCTIRYRQIGPLIDRQIFRFTTEG
MVNARFDVNDNSFRVTSMQAVINETPLIDLYRYDDVSGKTEQFGKFGVIYDINQI
ITTAVMTHTKHFDAYGRMKEVQYEIFRSLMWMVTQYDNMGRVVRVKELKVGYPYNTTR

YSYEYDADGQLQTVSINDKPLWRYSDNLGNLHLLSPGNSARLTPLRYDIRDIRTRLG
DVQYKMDDEGFLRQGGDIFEYNSAGLLIKAYNRAGSWSVRYRVDGLGRRVSSKSSHS
HHLOFFYADLTNPTKVTHLYNHSSEITSLYDYLQHLFAMELSSGDEFYIACDNIQT
PLAVFSGTGLMIKQILYATAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDV
LAGWTSPPDHELWKHLSSNVMPNLYMFKNNNPISNSQDIKCFMTDVSNSWLLTFFGQ
LHNVIPGYPKPDMDAMEPSYELIHTQMTQEWNDNSILGVQCEVQKQLKAFVTLERF
DQLYGSTITSCQAPKTKFSSGSVFGKVKFALKDGRVTVTDIISVANEDGRRVAAI
LNHAHYLENLHFTIDGVDTYFVKPSPSEGLAILGLSGRRRTLENGVNVTVSQINTV
LNGRTRRYTDIQLQYGALCLNTRYGTTLDEEKARVLELARQRAVRQAWAREQQRLREG
EEGLRAWTEGEKQQVLSTGRVQYDGFVFVISVEQYPELSDSANNIHFMRQSEMGRR"

ORIGIN

Alignment Scores:
Pred. No.: 3.9e-110 Length: 8624
Score: 1203.50 Matches: 234
Percent Similarity: 88.97% Conservative: 0
Best Local Similarity: 88.97% Mismatches: 1
Query Match: 86.96% Indels: 29
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x AB037723 (1-8624)

QY 27 LysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGlyTyrAlaLysAspAlaLys 46
|||||
Db 3 AAAAATGATGCCAACTGTGATTGTTTCTGGAGACGATGGTTATGCCAAGGATGCAAAG 62

QY 47 LeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGluLeuTyrValAlaAspLeu 66
|||||
Db 63 TTAATAACCCCATCTTCTTGGCTGTGTGCTGATGGGAGCTTACGTGGCCGACCTT 122

QY 67 GlyAsnIleArgIleArgPheIleArgLysAsnLysProPheLeuAsnThrGlnAsnMet 86
|||||
Db 123 GGGACATCCGAATTCGGTTTATCCGGAAGAACAGCCCTTCTCTCAACACCCAGAACATG 182

QY 87 TyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHis 106
|||||
Db 183 TATGAGCTGTCTTACCAATTGACCAGGAGCTCTATCTGTTTGATACCCCGGCAAGCAC 242

QY 107 LeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsnPheThrTyrThrGlyAsp 126
|||||
Db 243 CTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTCACTACACTGGGGAC 302

QY 127 GlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSer 146
|||||
Db 303 GCGACATCACACTCATCACAGACAACAATGGCAACATGGTAAATGTCCGCCGAGACTCT 362

QY 147 ThrGlyMetProLeuTrpLeuValValProAspGlyGlnValTyrTrpValThrMetGly 166
|||||
Db 363 ACTGGGATGCCCCCTCTGGCTGGTGTCCAGATGGCCAGGTGTACTGGGTGACCATGGGC 422

QY 167 ThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMetThrTyr 186
|||||
Db 423 ACCAACAGTGCACTCAAGAGTGTGACCACACAGGACAGAGTTGGCCATGATGACATAC 482

QY 187 HisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsnGlyTrpThrThrPheTyr 206
|||||
Db 483 CATGGCAATTCCGGCCCTCTGGCAACCAAGCAATGAAACCGATGGACCAACATTTTAT 542

QY 207 Glu----- 207

Db 543 GA-CAACCTCAGATGGCAGGAATTATCATCCACATTTTATCATATGAAGTACAGAGAGA 601

QY 208 -----TyrAspSerPheGlyArgLeuThrAsnValThr 218
|||||
Db 602 TTAGATAACAAGCCTGAAACACACACGGTACGACAGCTTGGCCGCTGACAAATGTGACC 661

QY 219 PheProThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGln 238
|||||
Db 662 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGTATACAGACAGTTTCAGTGCATGTCAG 721

QY 239 ValGluThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAla 258
|||||
Db 722 GTAGAGACCTCCAGCAAGGATGATGTACCATACCAACCAACCAACCTGTCTGCTCAGGCGCC 781

```
QY      259 PheTyrThr 261
Db      782 TTCTACACA 790

RESULT 8
AK127705
LOCUS
DEFINITION Homo sapiens cDNA FLJ45805 fis, clone NT2RI3007095, highly similar
to Mus musculus neuregulin 1 (Nrg1).
ACCESSION AK127705
VERSION AK127705.1 GI:34534730
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S.,
Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hotuta, T., Watanabe, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Takahashi-Fujii, A., Oshima, A., Sugiyama, A.,
Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K.
and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3048)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
    source
        location/Qualifiers
            1..3048
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="NT2RI3007095"
                /cell_line="NT2"
                /cell_type="teratocarcinoma"
                /clone_lib="NT2RI3"
                /note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
                precursor cells treated 2-weeks mitotic inhibitor after
                5-weeks retinoic acid (RA) induction.-majorly NT2 neuron"
ORIGIN
Alignment Scores:
Pred. No.: 2.47e-103 Length: 3048
Score: 1130.00 Matches: 214
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 81.65% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AK127705 (1-3048)

QY      48 AsnThrProSerSerLeuAlaValCysAlaAspGlyGluLeuTyrValAlaAspLeuGly 67
Db      1 AATACCCCATCTCTCTGGTGTGTGTGCTGATGGGAGCTCTACGTGGCGACCTTGGG 60

QY      68 AsnIleArgIleArgPheIleArgLysAsnLysPropheLeuAsnThrGlnAsnMetTyr 87
Db      61 AACATCCGAATTCGGTTTATCCGGAAGAACACAGCCTTTCTCTCAACACCCAGACATGTAT 120
```

```
QY      88 GluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLeu 107
Db      121 GAGCTGTCTTACCAATTGACAGGAGCTCTATCTGTTGATACCAACCGGCAAGCACCTG 180

QY      108 TyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsnPheThrTyrThrGlyAspGly 127
Db      181 TACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTACCTACACTGGGACGGC 240

QY      128 AspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSerThr 147
Db      241 GACATCACACTCATCACAGACAAACAATGGCAACATGGTAATGTCCGCCGAGACTTACT 300

QY      148 GlyMetProLeuTrpLeuValValProAspGlyGlnValTyrTrpValThrMetGlyThr 167
Db      301 GGGATGCCCTCTGGTGGTGGTCCCAGATGGCCAGGTGTTACTGGGTGACCATGGGACC 360

QY      168 AsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHis 187
Db      361 AACAGTGCACTCAAGAGTGTGACCAACACAGGACACAGAGTGGCCATGATGACATACCAT 420

QY      188 GlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsnGlyTrpThrPheTyrGlu 207
Db      421 GGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACAAACATTTTATGAG 480

QY      208 TyrAspSerPheGlyArgLeuThrAsnValThrPheProThrGlyGlnValSerSerPhe 227
Db      481 TACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCCTACCGCCAGGTGAGCAGTTTC 540

QY      228 ArgSerAspThrAspSerSerValHisValGlnValGluThrSerSerLysAspAspVal 247
Db      541 CGAAGTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAGACCTCCAGCAAGGATGATGC 600

QY      248 ThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyrThr 261
Db      601 ACCATAACCAACCAACCTGTCTGCTCAGGTGCCTTCTTACACA 642

RESULT 9
AP000716 112872 bp DNA linear HTG 14-SEP-2000
LOCUS Homo sapiens chromosome 11 clone CMB9-72M1 map 11q14, WORKING DRAFT
DEFINITION SEQUENCE, 8 unordered pieces.
ACCESSION AP000716
VERSION AP000716.3 GI:10129908
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 112872)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens 112,872 genomic DNA of 11q14
JOURNAL Published Only in DataBase (1999)
REFERENCE 2 (bases 1 to 112872)
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagami-hara, Kanagawa 228-8555, Japan
(E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-42-778-9923; Fax: 81-42-778-9924)
COMMENT On Sep 15, 2000 this sequence version replaced gi:8118885.

----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: CMB9-72M1
----- Summary Statistics
```


Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 111232 bases at least Q40
Consensus quality: 111769 bases at least Q30
Consensus quality: 112045 bases at least Q20
Insert size: 112172; sum-of-contigs
Quality coverage: 13.62x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 36854 contig of 36854 bp in length
36955 60659 contig of 23705 bp in length
60760 75582 contig of 14823 bp in length
75683 89456 contig of 13774 bp in length
89557 99398 contig of 9842 bp in length
99499 105224 contig of 5726 bp in length
105325 108843 contig of 3519 bp in length
108944 112872 contig of 3929 bp in length

NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 36854: contig of 36854 bp in length
36855 36954: gap of 100 bp
36955 60659: contig of 23705 bp in length
60660 60759: gap of 100 bp
60760 75582: contig of 14823 bp in length
75583 75682: gap of 100 bp
75683 89456: contig of 13774 bp in length
89457 89556: gap of 100 bp
89557 99398: contig of 9842 bp in length
99399 99498: gap of 100 bp
99499 105224: contig of 5726 bp in length
105225 105324: gap of 100 bp
105325 108843: contig of 3519 bp in length
108844 108943: gap of 100 bp
108944 112872: contig of 3929 bp in length.

FEATURES
source
1..112872
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q14"
/clone="CMB9-72M1"
1..36854
/note="assembly_fragment"
36955..60659
/note="assembly_fragment"
60760..75582
/note="assembly_fragment clone_end:SP6 vector_side:left"
75683..89456
/note="assembly_fragment clone_end:T7 vector_side:right"
89557..99398
/note="assembly_fragment"
99499..105224
/note="assembly_fragment"
105325..108843
/note="assembly_fragment"
108944..112872
/note="assembly_fragment"

ORIGIN

Alignment Scores:

Pred. No.: 1.52e-99 Length: 112872
Score: 1112.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.35% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AP000716 (1-112872)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 77607 AAGATCAACCGCATCAGGCAGGTCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 77666
QY 21 ProSerGlyCysAspCysLysAsnAlaAsnCysAspCysPheSerGlyAspGly 40
Db 77667 CCCAGTGGCTGTGACTGTAAATGATGCCAACTGTATTGTTTTCTGGAGACGATGGT 77726
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 77727 TATGCCAAGGATGCAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGAG 77786
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 77787 CTCACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGAAGAACAAGCCCTTTC 77846
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 77847 CTCACACCCAGAACATGTATGAGCTGTCTTCACCAATTGACCAGGAGCTCTATCTGTTT 77906
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 77907 GATACACCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 77966
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 77967 TTCACCTACACTGGGGACGGCGACATCACACTCATCAGACAACAATGGCAACATGGTA 78026
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db 78027 AATGTCGCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTG 78086
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 78087 TACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCACACACAGGACACGAG 78146
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 78147 TTGGCCATGATGACATACCATGGCAATTCCGGCCCTTCTGGCAACCAAAAGCAATGAAAAAC 78206
QY 201 GlyTrpThrThrPheTyrGlu 207
Db 78207 GGATGGACAACATTTTATGAG 78227

RESULT 10

AP001141

LOCUS

DEFINITION

Homo sapiens chromosome 11 clone RP11-646O20 map 11q14, WORKING

DRAFT SEQUENCE, 23 unordered pieces.

AP001141

ACCESSION

AP001141.2 GI:8118459

VERSION

HTG; HTGS_PHASE1; HTGS_DRAFT.

KEYWORDS

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 165,812 genomic DNA of 11q14

Published Only in DataBase (2000)

REFERENCE

AUTHORS

2 (bases 1 to 165812)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE
JOURNAL

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (08-FEB-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997819.

COMMENT

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: RP11-646O20
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 144380 bases at least Q40
Consensus quality: 155099 bases at least Q30
Consensus quality: 160622 bases at least Q20
Insert size: 163612; sum-of-contigs
Quality coverage: 4.18x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
23 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1 15120 contig of 15120 bp in length
15221 34407 contig of 19187 bp in length
34508 50417 contig of 15910 bp in length
50518 61510 contig of 10993 bp in length
61611 74361 contig of 12751 bp in length
74462 85865 contig of 11404 bp in length
85966 95879 contig of 9914 bp in length
95980 105221 contig of 9242 bp in length
105322 111606 contig of 6285 bp in length
111707 118475 contig of 6769 bp in length
118576 126493 contig of 7918 bp in length
126594 131617 contig of 5024 bp in length
131718 136431 contig of 4714 bp in length
136532 141754 contig of 5223 bp in length
141855 146576 contig of 4722 bp in length
146677 151463 contig of 4787 bp in length
151564 154735 contig of 3172 bp in length
154836 156703 contig of 1868 bp in length
156804 158461 contig of 1658 bp in length
158562 160497 contig of 1936 bp in length
160598 162229 contig of 1632 bp in length
162330 164116 contig of 1787 bp in length
164217 165812 contig of 1596 bp in length

Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 15120: contig of 15120 bp in length
* 15121 15220: gap of 100 bp
* 15221 34407: contig of 19187 bp in length
* 34408 34507: gap of 100 bp
* 34508 50417: contig of 15910 bp in length
* 50418 50517: gap of 100 bp
* 50518 61510: contig of 10993 bp in length

* 61511 61610: gap of 100 bp
* 61611 74361: contig of 12751 bp in length
* 74362 74461: gap of 100 bp
* 74462 85865: contig of 11404 bp in length
* 85866 85965: gap of 100 bp
* 85966 95879: contig of 9914 bp in length
* 95880 95979: gap of 100 bp
* 95980 105221: contig of 9242 bp in length
* 105222 105321: gap of 100 bp
* 105322 111606: contig of 6285 bp in length
* 111607 111706: gap of 100 bp
* 111707 118475: contig of 6769 bp in length
* 118476 118575: gap of 100 bp
* 118576 126493: contig of 7918 bp in length
* 126494 126593: gap of 100 bp
* 126594 131617: contig of 5024 bp in length
* 131618 131717: gap of 100 bp
* 131718 136431: contig of 4714 bp in length
* 136432 136531: gap of 100 bp
* 136532 141754: contig of 5223 bp in length
* 141755 141854: gap of 100 bp
* 141855 146576: contig of 4722 bp in length
* 146577 146676: gap of 100 bp
* 146677 151463: contig of 4787 bp in length
* 151464 151563: gap of 100 bp
* 151564 154735: contig of 3172 bp in length
* 154736 154835: gap of 100 bp
* 154836 156703: contig of 1868 bp in length
* 156704 156803: gap of 100 bp
* 156804 158461: contig of 1658 bp in length
* 158462 158561: gap of 100 bp
* 158562 160497: contig of 1936 bp in length
* 160498 160597: gap of 100 bp
* 160598 162229: contig of 1632 bp in length
* 162230 162329: gap of 100 bp
* 162330 164116: contig of 1787 bp in length
* 164117 164216: gap of 100 bp
* 164217 165812: contig of 1596 bp in length.

FEATURES

source

1. .165812
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q14"
/clone="RP11-646O20"

misc_feature

1. .15120
/note="assembly_fragment clone_end:SP6 vector_side:right"

misc_feature

15221. .34407
/note="assembly_fragment"

misc_feature

34508. .50417
/note="assembly_fragment"

misc_feature

50518. .61510
/note="assembly_fragment"

misc_feature

61611. .74361
/note="assembly_fragment"

misc_feature

74462. .85865
/note="assembly_fragment"

misc_feature

85966. .95879
/note="assembly_fragment"

misc_feature

95980. .105221
/note="assembly_fragment"

misc_feature

105322. .111606
/note="assembly_fragment"

misc_feature

111707. .118475
/note="assembly_fragment clone_end:T7 vector_side:right"

misc_feature

118576. .126493
/note="assembly_fragment"

misc_feature

126594. .131617
/note="assembly_fragment"

misc_feature

131718. .136431
/note="assembly_fragment"

misc_feature

136532. .141754


```
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 29785: contig of 29785 bp in length
* 29786 29885: gap of 100 bp
* 29886 52941: contig of 23056 bp in length
* 52942 53041: gap of 100 bp
* 53042 71728: contig of 18687 bp in length
* 71729 71828: gap of 100 bp
* 71829 87672: contig of 15844 bp in length
* 87673 87772: gap of 100 bp
* 87773 102073: contig of 14301 bp in length
* 102074 102173: gap of 100 bp
* 102174 109422: contig of 7249 bp in length
* 109423 109522: gap of 100 bp
* 109523 117422: contig of 7900 bp in length
* 117423 117522: gap of 100 bp
* 117523 124911: contig of 7389 bp in length
* 124912 125011: gap of 100 bp
* 125012 131158: contig of 6147 bp in length
* 131159 131258: gap of 100 bp
* 131259 135747: contig of 4489 bp in length
* 135748 135847: gap of 100 bp
* 135848 141547: contig of 5700 bp in length
* 141548 141647: gap of 100 bp
* 141648 147053: contig of 5406 bp in length
* 147054 147153: gap of 100 bp
* 147154 151529: contig of 4376 bp in length
* 151530 151629: gap of 100 bp
* 151630 156372: contig of 4743 bp in length
* 156373 156472: gap of 100 bp
* 156473 161167: contig of 4695 bp in length
* 161168 161267: gap of 100 bp
* 161268 164102: contig of 2835 bp in length
* 164103 164202: gap of 100 bp
* 164203 167603: contig of 3401 bp in length
* 167604 167703: gap of 100 bp
* 167704 170156: contig of 2453 bp in length.
*
FEATURES             Location/Qualifiers
  1..170156
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11q"
    /clone="RP11-7H7"
    /note="assembly_fragment"
  1..29785
    /note="assembly_fragment"
  29886..52941
    /note="assembly_fragment"
  53042..71728
    /note="assembly_fragment"
  71829..87672
    /note="assembly_fragment"
  87773..102073
    /note="assembly_fragment"
  102174..109422
    /note="assembly_fragment"
  109523..117422
    /note="assembly_fragment"
  117523..124911
    /note="assembly_fragment clone_end:T7 vector_side:left"
  125012..131158
    /note="assembly_fragment clone_end:SP6 vector_side:left"
  131259..135747
    /note="assembly_fragment"
  135848..141547
    /note="assembly_fragment"
  141648..147053
```

```
misc_feature          /note="assembly_fragment"
147154..151529
misc_feature          /note="assembly_fragment"
151630..156372
misc_feature          /note="assembly_fragment"
156473..161167
misc_feature          /note="assembly_fragment"
161268..164102
misc_feature          /note="assembly_fragment"
164203..167603
misc_feature          /note="assembly_fragment"
167704..170156
misc_feature          /note="assembly_fragment"

ORIGIN
Alignment Scores:      2.55e-99      Length:      170156
Pred. No.:            1112.00      Matches:      207
Score:                 100.00%      Conservative:  0
Percent Similarity:    100.00%      Mismatches:   0
Best Local Similarity: 100.00%      Indels:        0
Query Match:          80.35%        Gaps:          0
DB:                    2

US-10-029-020-14_COPY_1490_1750 (1-261) x AP002412 (1-170156)

QY      1  LysileAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
      |||
Db      143299 AAGATCAACCGCATCAGGCAGGTCACCACCTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 143358

QY      21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
      |||
Db      143359 CCCAGTGGCTGTGACTGTAAATATGATGCCAATCTTCTTGGCTGTGTGCTGATGGT 143418

QY      41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
      |||
Db      143419 TATGCCAAGGATGCAAGTTAAATATACCCCATCTTCTTGGCTGTGTGCTGATGGGAG 143478

QY      61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
      |||
Db      143479 CTCTACGTGGCGGACCTTGGGAACATCCGAATTCGGTTATCCGGAAGAACAAAGCCTTTC 143538

QY      81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuphe 100
      |||
Db      143539 CTCACACACCCAGAACATGTATGAGCTGTCTTACCAATTGACGAGACTATCTCTGTTT 143598

QY      101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
      |||
Db      143599 GATACCACCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 143658

QY      121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
      |||
Db      143659 TTCACCTACACTGGGGACGGCGACATCACATCATCACAGACAACAATGGCAACATGGTA 143718

QY      141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
      |||
Db      143719 AATGTCCGCCGAGACTCTACTGGGATGCCCCCTCTGGCTGGTGGTCCCAGATGCCAGGTG 143778

QY      161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
      |||
Db      143779 TACTGGGTGACCATGGGCACCAACAGTGCATCTCAAGAGTGTGACCACACAGGACACGAG 143838

QY      181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
      |||
Db      143839 TTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAAC 143898

QY      201 GlyTrpThrThrPheTyrGlu 207
      |||
Db      143899 GGATGGACAACATTTATGAG 143919

RESULT 12
AP002768/c
LOCUS
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-673F18,
```

complete sequence.
ACCESSION AP002768
VERSION AP002768.3 GI:13429926
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2000)
2 (bases 1 to 186084)

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT On Mar 21, 2001 this sequence version replaced gi:11071944.
FEATURES
source
1..186084
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-673F18"

ORIGIN

Alignment Scores:
Pred. No.: 2.86e-99 Length: 186084
Score: 1112.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 80.35% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AP002768 (1-186084)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 71689 AAGATCAACCGCATCAGGCAGGTACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 71630

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 71629 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGT 71570

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 71569 TATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGGAG 71510

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 71509 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCCGTTTATCCGGAAGAACCAAGCCTTTC 71450

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 71449 CTCAACACCCAGAACATGTATGAGCTGTCTTCACCAATTGACCAGGAGCTCTATCTGTTT 71390

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 71389 GATACCAACCGCAAGCAGCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAAC 71330

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 71329 TTCACCTACACTGGGAGCGCGACATCACACTCATCACAGACAACAATGGCAACATGGTA 71270

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 71319 GATACCAACCGCAAGCAGCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAAC 71330

Db 71269 AATGTCCGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTG 71210

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 71209 TACTGGGTGACCATGGGCACCAACAGTGCCTCAAGAGTGTGACCACACAAGGACACGAG 71150

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 71149 TTGGCCATGATGATACCATCCATGGCAATTCGGCTTCTGGCAACCAAAAGCAATGAAAC 71090

QY 201 GlyTrpThrThrPheTyrGlu 207
Db 71089 GGATGGACAACATTTATGAG 71069

RESULT 13
AB026980
LOCUS AB026980 9264 bp mRNA linear VRT 05-NOV-1999
DEFINITION Danio rerio mRNA for ten-m4, complete cds.
ACCESSION AB026980
VERSION AB026980.1 GI:5307784
KEYWORDS ten-m4.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE
AUTHORS Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.
TITLE Compartmentalized expression of zebrafish ten-m3 and ten-m4, homologues of the Drosophila ten(m)/odd Oz gene, in the central nervous system
JOURNAL Mech. Dev. 87 (1-2), 223-227 (1999)
MEDLINE 99425191
PUBMED 10495292
REFERENCE 2 (bases 1 to 9264)
AUTHORS Mieda,M.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute, RIKEN, lab. for Developmental Gene Regulation; 2-1 Hiroosawa, Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp, Tel:81-48-467-9713, Fax:81-48-467-9714)

COMMENT Sequence updated (29-Jun-1999).
FEATURES
source
1..9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1..9264
/gene="ten-m4"
323..8797
/gene="ten-m4"
/note="similar to Drosophila melanogaster tenm/odz and Mus musculus Doc4"
/codon_start=1
/product="ten-m4"
/protein_id="BAA81893.1"
/db_xref="GI:5307785"
/translation="MEVKRRPYRSLTSSADSEDGKINPKSYSSSET LKAFDQDSRLAYGSRVKDLVHHEADFSRQGPDPFSLRDMAFGDPVPHMGAYRTEMGL PHRDYSVSVASDADTETDIMSPEHAEVRLWGRSNTKSGRSSCLFSRANSNLTLDTEH ENTENGPLHCSSASSSPVDSPPPPSHAAANQSQRLLGNSGAQGRDSESEDEFQPN SFLVKTSGNVCTPAAATANEGSFONHSRLRTPPLPLFHSHPHQHTASIGSLSRSN YTORSNPAPATDSSAPNEGPTSAQDSSSAQDNWLLNSNVPLETRNIAKQTFLETLDQ NFIEDILATARRDGAYTDGHFLFKPGGTSPLYCTTSPGYPLTSTSVSPPPRPLPRN TFSRPAFSLKKPYKCNWKAALSAILISVTLVFLLAYFIAMHLFGLNHLQPVQRQI YQLTEDNTSGLHLPDGLPLGNLTGLEFPDRGSRDDKLDGFFEDSFIDMGEIDVG RKVAQLIPPGIFWRQSVFIDHPMYLKFNVSLSKDALVGIYGRGLPPSHTQDFVELL DGRRLLSQGLPGLDGPFPFAQQRSILVPIITSHDTGCIQYMDSGIWHLAVYNDGKETEY SFLTATIDSIDDCPSNCFGNGDCVSGNCHCFPGGRCFPCGRASCPVLCSGNGQYLYKGR CMHSGWKSGSECDVPTNQCIDITCSGHGTCTIVGTICINPSYKGENCEEVDCLDPTCSG RGVCVRGECHCFVGGPGCESPRASCMEQCSGHGSFLADTNTCNDHNWTGDCSTE LCAADCGGHGICVAGSCRCDEGMWGTGCEQRACHPRCSEHGTCKDKGCECSPGWNGEH

CTIEGCPGLCNGNGRCTLGNGWYCVCOLGWRGAGCDTSMETACSDGKNDGDGLTDC
MDPCCQLQASCHTSLCVGSPDPLDIQETQISSSLSTLQSFYQRIHFLVGRDSTHVI
PDVNPFDGIHACVIRGOVWTSDDTGLVGNISFINPKPAYGTTITRODGSFDLVNSGGV
AIGLRFERAPFITQEHITLWLPWGRFFVMDTIVMRHEVNDIPSCDLSFTRPMPILVPA
PLTAFAGTCBERGIVVPEIQTLQEEVRIPGTDMLRGLYLSRSTSGYKSLRLITLTHSTI
PFLMKVHLMVAVEGRFLFRKWFSAAPNLISYDFWMDKTDVYSQKVYGLSEAFVSVGFY
ESCPDLILWEKRTAVLQYETTASNLGWSVDKHALNIQSGILHKGNGENIFISQOP
PVIISIMNGRRRSISCPSCNGLADGNKLLAPVALACGSDGLYGVDFNVVRRITFTG
NVSLELSNSPAHKYILATSPVSGWLYLSDTSRKYKFKVKSLYAVKDVAKNLELVAG
TGDCQLPYDETRCGDGGKAVEATLTNPRGITVDKYGVIFFVDGTMIRRIDQNGIISTL
LGFNDLTLSARPLSCDSVMDISQVRLEWPTDLAVSPMDNSLYLDNNVVLQISENHQVR
IVAGRPMHCQVPLDHFVLSKIAIHATLESANALAVSHNGLLYIAESDEKKINRVRQV
STNGEISLLAGAPSGCDCKNDANDCYSGDDGYAKDAKLNAPSSLAVSPDGELFIADL
SGTNPTISIKRNVTLPIDNGLNLVEWRQKEARGQVTVYGRRLRVHNRNLLSMDFDR
VTRTEKVYDDHRKFTLRIDYDHAGRPTLWAPSRNLGNVNTYSPGGHIAQIGRTMSV
RMEYDQNGRITSKIFADGKWSYTYLEKSMVLLYSQROYIFEEDKNDRLSVTPMPNV
ARQLETTRSIGYVRNTYRPEGNATVLQDYSDDGLLLQTIHQGTGRRVYKYGKLSR
LLEILYDTTRIAFSYDESAGMLKTGLQSEGFACITRYRQIGPLIDRQIFRPFSEGMV
NARFDYNDNSFRVTSMAVINETPLPIDLYRYDDVSGKTEQKFGFVIVYDINQIIT
TAVMTHTKHFDAYGRVKEVYEIIFRSLMYMMQFDNMGRVVAKELKVGVPYANTRYA
YEYDADGQLQVVSINDKPLWYSYDLNGLNHLSPNSARLTPLRYDIRDRITRLGDV
QYRLDEDFLQRGNDFPEYNSAGLLVKTYNKNGWTKIKRYDGLKVRNGVSSRSRSTQGH
LQFFYADLSSPTRVTHMYNHSSEITSLYDLOGLHFAFAMELSSGDEFYVACDNIGTPL
AVFSGAGLMIKQILHTAFGEVYLDNSPFLVYDQGLYEPLTKLVHMGRRDYDVLVA
GRWTPPDHDIRKRLSDNIVPFNLYMFKNNPLSNSQETKCYMTDVNSWLVTFGFQLY
NVIPGYRKPVTDAMEPSYELVHTQIKTOEWDSTKSVLGVCEVQRLKSFVRLERFGQ
IYSASDSGCPPTPLHLTAFATGTSLFGKGVKVAIREGRVEADIISLANEDGRRIRAAVLG
KASYLDLHLFTIAGLDTHYFVKSGLVEGDLSSLGTMVQRTLETGVNVTVSQVNMVLG
GRSRRITDIOMQYGTLSLNVRYGSSVDEEKVRVLELARQAVATAWAHERHLRQGE
GSAWTDGERQQLSSGRVQYEGFYIVSDQFPPELTDNINNHFWRQTEMGRR"

ORIGIN

Alignment Scores:
Pred. No.: 5.15e-100 Length: 9264
Score: 1103.00 Matches: 195
Percent Similarity: 90.04% Conservative: 40
Best Local Similarity: 74.71% Mismatches: 26
Query Match: 79.70% Indels: 0
DB: 5 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AB026980 (1-9264)

QY	1	LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla	20
Db	4955	AAAAATCAACCGTGTTCGTACAGGTTTCGACCAACGGTGAAATCTCTCTGCTCGCTGCC	5014
QY	21	ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly	40
Db	5015	CCCACTGGCTGCGACTGTAAAAATGATGCCAACTGTGACTGTCTACTCTGGCGACGCGC	5074
QY	41	TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu	60
Db	5075	TACGCGAAAGACGCCAAACTTAATGCTCCTTCATCGTTAGCGGTTTCGCTGACGGTGAG	5134
QY	61	LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe	80
Db	5135	CTTTTATCGCCGATTTAGGAACATTCGTATTCGATATGTTTCGTAGGAACAAAGCGTTC	5194
QY	81	LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe	100
Db	5195	CTAAATCCGCTCAATATGATGAGATTTCATCACCATTGACGACGAGCTTTACTTGT	5254
QY	101	AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn	120
Db	5255	GATGTAATCGGAGCCACGCTTTCACCCAAAGTTTGACCAACGGCGATTACCTTACAAC	5314
QY	121	PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal	140
Db	5315	TTCACATACTCCGGCGAGGGTGACTTGAGCAGCATCACTGATAAGATAAGAACCGGTG	5374

QY	141	AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal	160
Db	5375	AGCATCCGACAGAGACTCCACAGGCTGCCACTCTGGTTAATGGGCCCGATGACAGACC	5434
QY	161	TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu	180
Db	5435	TTCTGTTTCCACCATGGGACTAATAATCGCTGAAGAGCGTGGCAGCACAAAGGACAGAA	5494
QY	181	LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn	200
Db	5495	ATCGCCGTCATGACGCTACCATGGCAGCTCGGGCCTGCTGGCCACCAAGAGCAATGAGGAC	5554
QY	201	GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro	220
Db	5555	GGCTGGAGCACCTTCTATGATATGACAACTACGGCCGGCTGACAAACGTGACCTACCT	5614
QY	221	ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu	240
Db	5615	ACAGGCCGGGTGAGCAGTTATCGCACAGACTCAGACAGCAGCTGTACGGTGCACAAACCGAA	5674
QY	241	ThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr	260
Db	5675	GGCTCAACAAGAGGACATCATCCGTCAACCAACCTGTGACGATCCGGAACATTCTAC	5734
QY	261	Thr	261
Db	5735	ACG	5737

RESULT 14
AC073599/c

LOCUS AC073599 226256 bp DNA linear ROD 28-FEB-2001
DEFINITION Mus musculus 7 BAC RP23-8M3 (Roswell Park Cancer Institute Mouse BAC Library) complete sequence.

ACCESSION AC073599

VERSION AC073599.16 GI:13162472

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 226256)

REFERENCE

AUTHORS

Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C., Dederich,D., Thomas,S., Okwuonu,G., Carlock,C., Garner,T., Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J., Buhay,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z., Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S., Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R., Gorrell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogues,M., Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S., Kovar,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Martin,R., Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morris,S., Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Oguh,M., Parish,B., Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,L., Watlington,S., Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A., Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstock,G., Worley,K. and Gibbs,R.

Direct Submission

TITLE

JOURNAL

REFERENCE 2 (bases 1 to 226256)

AUTHORS Worley,K.C.

Direct Submission

JOURNAL

Submitted (26-JUN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 226256)

AUTHORS Worley,K.C.

Direct Submission

TITLE

JOURNAL

Submitted (28-FEB-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Feb 28, 2001 this sequence version replaced gi:12963002. Sequencing is completed to a minimum standard of double strand

coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES

source	Location/Qualifiers
	1. .226256
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="7"
	/clone="RP23-8M3"
misc_feature	complement(147. .670)
	/note="Region similar to Mm#S1670873 601761001F1 Mus musculus cDNA: BF124644"
repeat_region	938. .971
	/rpt_family="B1-F"
repeat_region	1118. .1165
	/rpt_family="T-rich"
repeat_region	complement(2209. .2323)
	/rpt_family="L2"
repeat_region	2468. .2493
	/rpt_family="CAAA)n"
repeat_region	complement(2497. .2613)
	/rpt_family="B1-F"
repeat_region	3397. .3501
	/rpt_family="MIR"
repeat_region	3758. .3948
	/rpt_family="B4A"
repeat_region	5668. .5732
	/rpt_family="B4"
repeat_region	5733. .5809
	/rpt_family="PB1D10"
repeat_region	complement(5737. .5848)
	/rpt_family="YREP_MM"
repeat_region	5810. .5820
	/rpt_family="B4"
repeat_region	6217. .6254
	/rpt_family="TTGG)n"
repeat_region	6509. .6566
	/rpt_family="TG)n"
repeat_region	6721. .6756
	/rpt_family="CA)n"
repeat_region	complement(7827. .7942)
	/rpt_family="PB1D10"
repeat_region	complement(7899. .7964)
	/rpt_family="B4"
repeat_region	complement(7965. .8038)
	/rpt_family="ID4"
repeat_region	8093. .8191
	/rpt_family="L2"
repeat_region	8326. .8437
	/rpt_family="MIR"
repeat_region	9579. .9616
	/rpt_family="CATA)n"
repeat_region	9641. .9664
	/rpt_family="CA)n"
repeat_region	complement(9715. .9783)
	/rpt_family="PB1D10"
repeat_region	complement(9778. .9830)
	/rpt_family="ID3"
repeat_region	9927. .9958
	/rpt_family="CAAA)n"
repeat_region	complement(10121. .10237)

repeat_region	/rpt_family="B1-F"
	10243. .10414
	/rpt_family="CCTG)n"
repeat_region	10440. .10509
	/rpt_family="CT-rich"
repeat_region	complement(10532. .10578)
	/rpt_family="B1-F"
repeat_region	10971. .11355
	/rpt_family="Lx6"
misc_feature	complement(11571. .13187)
	/note="Region similar to Mm#S685488 Odd Oz/ten-m homolog 4 (Drosophila): AB025413"
repeat_region	14505. .14614
	/rpt_family="PB1"
repeat_region	15111. .15194
	/rpt_family="MIR"
repeat_region	complement(15863. .16043)
	/rpt_family="B4"
repeat_region	16385. .16459
	/rpt_family="MIR"
repeat_region	19475. .19693
	/rpt_family="B2"
repeat_region	20006. .20169
	/rpt_family="B3"
repeat_region	20170. .20190
	/rpt_family="(TA)n"
repeat_region	20191. .20251
	/rpt_family="B3"
repeat_region	complement(21402. .21484)
	/rpt_family="MIR"
repeat_region	24484. .24563
	/rpt_family="L2"
repeat_region	24789. .24907
	/rpt_family="L2"
repeat_region	complement(25094. .25241)
	/rpt_family="MIR"
repeat_region	complement(26937. .27347)
	/rpt_family="LIMCc"
repeat_region	27569. .27613
	/rpt_family="(TTTA)n"
repeat_region	complement(28594. .28633)
	/rpt_family="L2"
repeat_region	complement(28841. .28955)
	/rpt_family="PB1D10"
repeat_region	29428. .29480
	/rpt_family="(TATG)n"
repeat_region	29510. .29582
	/rpt_family="CAGA)n"
repeat_region	29583. .29723
	/rpt_family="GA-rich"
repeat_region	29776. .29959
	/rpt_family="(CAGAGA)n"
repeat_region	complement(30193. .30338)
	/rpt_family="B1_MM"
repeat_region	31162. .31452
	/rpt_family="B4"
repeat_region	31913. .31950
	/rpt_family="AT-rich"
repeat_region	32864. .32885
	/rpt_family="AT-rich"
repeat_region	complement(33023. .33070)
	/rpt_family="MIR"
repeat_region	34393. .34458
	/rpt_family="MIR"
repeat_region	complement(34687. .34859)
	/rpt_family="MER5B"
repeat_region	34872. .34950
	/rpt_family="MER5A"
repeat_region	35660. .35869
	/rpt_family="URR1A"
repeat_region	36089. .36106
	/rpt_family="B4"
repeat_region	36107. .36239

/rpt_family="B1_MM"	
repeat_region	36263. .36295
/rpt_family="(CCA)n"	
repeat_region	36398. .36419
/rpt_family="(CCCCAG)n"	
repeat_region	36790. .36817
/rpt_family="(T)n"	
repeat_region	complement(38755. .38926)
/rpt_family="B4"	
repeat_region	complement(39294. .39403)
/rpt_family="L2"	
repeat_region	40007. .40038

Alignment Scores:			
Pred. No.:	9.29e-98	Length:	226256
Score:	1098.00	Matches:	204
Percent Similarity:	99.52%	Conservative:	2
Best Local Similarity:	98.55%	Mismatches:	1
Query Match:	79.34%	Indels:	0
DB:	10	Gaps:	0

US-10-029-020-14_COPY_1490_1750 (1-261) x AC073599 (1-226256)			
QY	1	LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla	20
Db	33926	AAGATCAATCGCATCAGGCAGGTCACTACAAGTGGTGAGATCTCACTGGTTGCTGGTGCC	33867
QY	21	ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly	40
Db	33866	CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGACTGCTTCTCTGGAGATGATGGT	33807
QY	41	TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu	60
Db	33806	TACGCCAAGGATGCAAAAGCTGAATACCCCATCGTCTTGGCTGTGTGTGCTGACGGGAG	33747
QY	61	LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe	80
Db	33746	CTCTATGTGGCCGACCTGGGAAACATCCGAATTCGATTATCCGGAGAACAAGCCTTTC	33687
QY	81	LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe	100
Db	33686	CTGAACACTCAGAACATGTACGAGCTATCTCCCCCATCGACCGAGGAGCTGTACCTCTTT	33627
QY	101	AspThrTyrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn	120
Db	33626	GATACCAAGTGGCAAGCATCTGTACACTCAGAGCCTACCCACAGGGGACTACCTGTACAAC	33567
QY	121	PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal	140
Db	33566	TTCACCTACACAGGGGACGGGACATCACACATATCACCGACAACAATGGCAACATGGTG	33507
QY	141	AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGlnVal	160
Db	33506	AACGTCCGCCGAGACTCTACCGGGATGCCTCTCTGGCTGGTAGTCCAGATGGCCAGGTA	33447
QY	161	TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu	180
Db	33446	TACTGGGTAAACCATGGGCACCAACAGCGCACTCAGAAGTGTGACCACACACAGGACCGAG	33387
QY	181	LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn	200
Db	33386	CTAGCCATGATGACCTACCATGGCAACTCTGGCCTCTTGGCAACCAAAAGCAATGAAAC	33327
QY	201	GlyTrpThrThrPheTyrGlu	207
Db	33326	GGGTGGACAACGTTTTATGAG	33306

RESULT 15			
AC120288/c			
LOCUS	AC120288	245722 bp	DNA linear HTG 15-NOV-2002
DEFINITION	Rattus norvegicus clone CH230-17M17, WORKING DRAFT SEQUENCE.		
ACCESSION	AC120288		
VERSION	AC120288.4 GI:25008439		

KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE	Rattus norvegicus (Norway rat)
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE	1 (bases 1 to 245722)
AUTHORS	Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D., Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L., Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D., Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W., Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 245722)
AUTHORS	Worley,K.C.
TITLE	Direct Submission
JOURNAL	Submitted (05-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 245722)
AUTHORS	Rat Genome Sequencing Consortium.
TITLE	Direct Submission
JOURNAL	Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	On Nov 15, 2002 this sequence version replaced gi:23669170. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas

(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GWUK
Center clone name: CH230-17M17
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 234443 bases at least Q40
Consensus quality: 234913 bases at least Q30
Consensus quality: 236024 bases at least Q20
Estimated insert size: 236455; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 245722: contig of 245722 bp in length.

FEATURES
source

1. .245722
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-17M17"
1. .1076
/note="wgs_end_extension
clone_end:Sp6"
6381. .7227
/note="clone_boundary
clone_end:Sp6"
site:
end_sequence:BH280179"
6388. .6810
/note="clone_boundary
clone_end:Sp6"
site:
end_sequence:BH280179"
complement(73338..240671)
/note="clone_boundary
clone_end:T7"
site:
end_sequence:BH280178"
complement(240061..240644)
/note="clone_boundary
clone_end:T7"
site:
end_sequence:BH280178"
244240. .245722
/note="wgs_end_extension
clone_end:Sp6"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Alignment Scores:

Pred. No.:	5.19e-97	Length:	245722
Score:	1091.00	Matches:	203
Percent Similarity:	99.03%	Conservative:	2
Best Local Similarity:	98.07%	Mismatches:	2
Query Match:	78.83%	Indels:	0
DB:	2	Gaps:	0
US-10-029-020-14_COPY_1490_1750 (1-261) x AC120288 (1-245722)			
QY	1	LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla	20
Db	15709	AAGATCAACCGTATCAGGCAGGTACGACAAAGTGTGAGATCTCACTGGTTGCTGTGCC	15650
QY	21	ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly	40
Db	15649	CCCAGCGGCTGCGACTGTAAAAACGACGCCCACTCGACTGCTTCTCTGGAGATGATGGC	15590
QY	41	TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaaspGlyGlu	60
Db	15589	TACGCCAAGGACGCAAAAGCTGAACACCCCATCTTCCTTGGCCGTGTGTGCAGACGGGAG	15530
QY	61	LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe	80
Db	15529	CTCTACGTGGCCGACCTTGGAAACATCCGGATCCGGTTTATCCGGAAGAACAAAGCCTGTC	15470
QY	81	LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe	100
Db	15469	CTGAACACCCAGAACATGTATGAGCTGTCTCCCCCATCGACCAGGAGCTGTACCTCTTT	15410
QY	101	AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn	120
Db	15409	GACACCAGTGGCAAGCATCTGTACACTCAGAGCCTGCCACAGGAGACTACCTGTACAAC	15350
QY	121	PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal	140
Db	15349	TTCACTTACACGGGGGACGGGGACATCACACACATCACAGACAACATGGCAACATGGTG	15290
QY	141	AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGlnVal	160
Db	15289	AACGTCCGCCGAGATTCTACCGGGATGCCTCTCTGGCTGGTAGTCCCAGATGGCCAGTG	15230
QY	161	TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu	180
Db	15229	TACTGGTAACCATGGGCACCAACAGCGCACTCAGAAAGTGTGACCACGCAAGGACACGAG	15170
QY	181	LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn	200
Db	15169	CTGGCTATGATGACATACCATGCCAACTCTGGCCTCTTGGCGACCAAAAGCAATGAAAAAC	15110
QY	201	GlyTrpThrThrPheTyrGlu	207
Db	15109	GGGTGACAAACGTTTTTATGAG	15089

Search completed: August 14, 2004, 12:00:37
Job time : 5091.26 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 1786.37 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200
Perfect score: 549
Sequence: 1 RLFRKWFAAAPDLSYVFTWD.....IQSGILHKGNGENQFVSQQP 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_in: *
18: em_mu: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_un: *

29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pln: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	549	100.0	8354	6	AX556500	AX556500 Sequence
2	549	100.0	8645	6	AX600210	AX600210 Sequence
3	544	99.1	9722	10	AF059485	AF059485 Mus muscu
4	540	98.4	8585	10	AB025413	AB025413 Mus muscu
5	538.5	98.1	8438	6	AX675551	AX675551 Sequence
6	484	88.2	9264	5	AB026980	AB026980 Danio rer
7	454	82.7	5804	10	AK122513	AK122513 Mus muscu
8	454	82.7	7816	10	AF195418	AF195418 Mus muscu
9	454	82.7	8964	10	AB025412	AB025412 Mus muscu
10	453	82.5	3394	9	AK125869	AK125869 Homo sapi
11	453	82.5	5309	9	AB040888	AB040888 Homo sapi
12	453	82.5	8473	6	AX662357	AX662357 Sequence
13	453	82.5	8487	6	AX662359	AX662359 Sequence
14	453	82.5	8645	6	AX662355	AX662355 Sequence
15	453	82.5	8675	6	AX662353	AX662353 Sequence
16	446	81.2	8816	5	AB026979	AB026979 Danio rer
17	426	77.6	8409	5	GGA279031	AJ279031 Gallus ga
18	426	77.6	8409	6	AX250068	AX250068 Sequence
19	420	76.5	8118	5	GGA238613	AJ238613 Gallus ga
20	417	76.0	8797	6	AX250063	AX250063 Sequence
21	417	76.0	8797	6	AX250066	AX250066 Sequence
22	417	76.0	8797	10	AB025411	AB025411 Mus muscu
23	416	75.8	6560	6	AX250065	AX250065 Sequence
24	416	75.8	7713	9	HSM808325	BX648178 Homo sapi
25	416	75.8	7781	9	AB032953	AB032953 Homo sapi
26	411	74.9	8297	9	AF100772	AF100772 Homo sapi
27	408	74.3	8373	10	AB025410	AB025410 Mus muscu
28	407	74.1	9729	6	AX250013	AX250013 Sequence
29	407	74.1	9826	6	AX250008	AX250008 Sequence
30	406	74.0	8689	6	AX250067	AX250067 Sequence
31	406	74.0	8689	10	AF086607	AF086607 Rattus no
32	323	58.8	8575	6	AX921803	AX921803 Sequence
33	265	48.3	112872	2	AP000716	AP000716 Homo sapi
34	265	48.3	170156	2	AP002412	AP002412 Homo sapi
35	265	48.3	186084	9	AP002768	AP002768 Homo sapi
36	261	47.5	226256	10	AC073599	AC073599 Mus muscu
37	261	47.5	245722	2	AC120288	AC120288 Rattus no
38	261	47.5	257693	2	AC103323	AC103323 Rattus no
39	260	47.4	180659	5	BX001038	BX001038 Zebrafish
40	248.5	45.3	10471	3	AF008228	AF008228 Drosophil
41	248.5	45.3	11451	3	DMTENN	X73154 D.melanogas
42	248.5	45.3	14045	3	DMODZ4	AF008227 Drosophil
43	248.5	45.3	82852	2	AC017685	AC017685 Drosophil
44	248.5	45.3	169841	3	AC009366	AC009366 Drosophil
45	248.5	45.3	305612	3	AE003597	AE003597 Drosophil

ALIGNMENTS

AX556500
LOCUS AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M., Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E., Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L., Smithson,G., Li,L. and Ji,W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source
1. .8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.2e-58 Length: 8354
Score: 549.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AX556500 (1-8354)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3332 CGCCTCTTCAGGAAGTGGTTTCGCTGCAGCCCCCAGACCTGTCTATTATTTTCATTGGGAC 3391
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3392 AAGACAGACGCTCTACAAACAGAAAGGTGTTTGGGCTTCAGAAAGCCTTTGTTTCCTGGGT 3451
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3452 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGCTTCAGAAAGCCTTTGTTTCCTGGGT 3511
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3512 GGCTATGAATTTGACGCGTCCAAAGCTTGGAGGATGGAGCTAGACAAACATCATGCGCTC 3571
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3572 AACATTCAAAGTGGTATCTCTGCACAAAGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAG 3631
QY 101 Pro 101
Db 3632 CCT 3634
RESULT 2
AX600210
LOCUS AX600210 8645 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 22 from Patent WO02072830.
ACCESSION AX600210
VERSION AX600210.1 GI:28400252
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S., Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G.,

Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K., Khare,R. and Walia,N.K.
TITLE Proteins associated with cell growth, differentiation, and death
JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7488573CB1"
ORIGIN
Alignment Scores:
Pred. No.: 1.25e-58 Length: 8645
Score: 549.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AX600210 (1-8645)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3402 CGCCTCTTCAGGAAGTGGTTTCGCTGCAGCCCCCAGACCTGTCTATTATTTTCATTGGGAC 3461
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3462 AAGACAGACGCTCTACAAACAGAAAGGTGTTTGGGCTTCAGAAAGCCTTTGTTTCCTGGGT 3521
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3522 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGTCTGCAG 3581
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3582 GGCTATGAATTTGATCGTCCAAAGCTTGGAGGATGGAGCTAGACAAACATCATGCGCTC 3641
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3642 AACATTCAAAGTGGCATCCTGCACAAAGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAG 3701
QY 101 Pro 101
Db 3702 CCT 3704
RESULT 3
AF059485
LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998
DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 9722)
AUTHORS Wang,X.Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P., Zinszner,H. and Ron,D.
TITLE Identification of novel stress-induced genes downstream of chop
JOURNAL EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE 98315054
PUBMED 9649432
REFERENCE 2 (bases 1 to 9722)
AUTHORS Wang,X.-Z. and Ron,D.
TITLE Direct Submission
JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University Medical Center, 550 First Ave., New York, NY 10016, USA
FEATURES
source
1. .9722
/organism="Mus musculus"


```
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1. .9722
/gene="Doc4"
/note="downstream of CHOP 4"
583. .9060
/gene="Doc4"
/note="similar to Drosophila melanogaster tenm/odz and human gamma-heretulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
translation="MDVKERKPYRSLTFRRDAERRYTSSSADSEEGKGPQKSYSSSET
LKAYDQARLAYGSRVKDMVPQEAEEFCRTGTNFTLRELGLGEMTPPHGTLYRTDIGL
PHCGYSMGASSDADLEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTEHE
NTETGAPLHCSSASTPIEQSPSPPPPANESQRRLLGNVQAQPTPDSDEEEFVFN
SFLVSGSASLGAANDHPSSLQNHPRLKTPPPLPHAHTPNQHHASINSLNRGNFT
PRSNPSAPATDLSLGEPPAGSAQEPHAQDNVNSKIPVETRNLGKQPFPLGTWQDN
LIEMDIPASARRDGAYSDGHFFKPGTSPLEFCTTSPGYPLTSTSYVSPPPRLPRST
FSRPAFLKKPSKYCNWKAALISATLVILLAYFVAMHLFGLNWHLQPMEGQM
MYEITEDASSWPVPTDVSYPGGTGLETPDRKGAAEGKPSLFPEDSFIDSGEI
DVGRASQKIPPGTFWRFSQVFIHPVHLKFNVSGLKAALVGIYGRKGLPPSHTQLDFV
ELLDGRLLTQEARSLLEGPQRQSRGPVPPSSHETGFIQYLDSGIWHLAFYNDGKESEV
VSFLTALIESVDNCPNCGDCTSGTCHFCFLGFLGPDGCRASCPVLCSGNGQYMKG
RCLCHSGWKAEDVPTNQCIDVACSSHGTCIMGTICINPGYKGESCEEVDGMDPTCS
SRGVCVRGECHCSVGWGTNCETPRATCLDQCSGHGTFPLDTGLCNDPSSWTGHDCSI
EICAADCGHGVCVGGTCRCEDEGMWGAACQDRAQCHPRCAEHGTCRDGKCECTPGWNGE
HCTIBGCPGLCNGNRCITLDLNGHVCVQLGWRGTGCDTSMETCGDGKDNDDGLVD
CMDPDCCLQPLCHVNPCLGSPDLDIIQETAPVSQQNLNSFYDRIKFLVGRDSTHS
IPGENPFDGGHACVIRGOVMTSDGTPLVGNISFINNPLFGYTI SRQDGSFDLVNNG
ISILIRFERAPFITQEHTLWLPWDRFFVMEITIVMRHEENETPSRDLNSFARNPVVSP
SPLTSPASSCAEKPIVPEIQALQEEI VIAGCKMRLSYLSRTPGYKSVVRISLTHPT
IPFNLMKVHLMVAVEGRLFRKWFAPAAPDLISYFIWDKTDVYNQKVFGLSEAFVSGYE
YESCPDLILWEKRTAVLQGYEIDASKLGGWSLDKHHALNQSGILHKGNGENQFVSQ
PVIQSIMGNRRRSISCPSCNGLADGNKLLAPVALTCSGDSGLVGVDFNYIRIFFS
GNVTNILEMHSHPAHKYI LATDPMGAVFLSDTNRRVFKVKTSTTVVKDLVKNSVVA
GTGDQCLPFPDTRCGDGGKATEATLTNPRGITVDKFGLLYFVDGTMIRRVQDNGLIST
LLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYLDNNVLOISENHQV
RIVAGRPMHCQVPGIDQFLLSKVAIHATLESATALAVSHNGVLVIAETDEKKINRIQ
VTTSGEISLVAGAPSGCDKNDANCDFSGDDGAKDAKLNTPSSLAVCADGELYVAD
LGNIRIRFIRKNKPFLNTQMYELSSPIDQELYLFTDSGKHLYTQSLPTGXYLNFTY
TGDGDITHITDNNGNMNVRRDSTGMPLWLVPDGVVYMTGMTNSALRSVTTQGH
AMMTYHNSGLLATKSNENGWTFEYDSFGRLLTNVTFPTQGVSSFRSDTDSVHVQV
ETSSKDDVTITTHLSGSAFYLLQDQVRNSYIIGADGSLRLLLANGMEVALQTEPHL
LAGTVNPTVTKRNVTLPIDNGLNLVWRQKEQAGQVTVFGPRLRVHNRNLLSLDFD
RVTRTEKIYDDHRKFTLRILYDQAGRPSPFWSPSSRLNGVNVTVSPGGHIAGIQRGIMS
ERMEYDQAGRITSRIFADGKWSYTYLEKSMVLHLHSQRQYIFEFDKNDRLSSVTMPN
VARQTLLETIRSVGYRNIYQPPEGNASVIQDFTEDHLLHTFYLGTGRRVIYKYGKLS
KLAETLYDTTKVSFTYDETA GMLKTVNLQNEGFTCTIRYQIGPLIDRQIFRFTEEGM
VNARFDYNDNSFRVTSMAVINETPLPIDLYRVDVSGKTEQFGKFGVIYDINQII
TTAVMTHSKHFDAYGRMKEVQYETFRSLMYMTVQYDNMGRVVVKELKVGVPYANTRY
VQYKMEDEGSLRQGGVFEYNSAGLLIKAYNRASGWSVRYRYDGLGRRVSSKSSHSH
SYEYDADGQLQTVSINDKPLWRYSYDLNGLNLLSPGNSARLTPLRYDLDRITRLGD
HLQFFFYADLTNPTKVTHLYSHSSSEITSLYIDLQGHLPAMELSSGDEFYIACDNIGTP
LAVFSGTGLMIKQIITYAYGEIYMDTNPNFQIIGYHGGLYDPLTKLVHMGRRDYDL
AGRWTSPDHELWKLSSNSIVPFHLYMFKNPNPISNQDIKCFMTDVSMLLTFGFOL
HNVIPGYPKPDTDAMEPSYELVHTQMTQEWDNKSILGVQCEVQKQKAFVTLERFD
QLYGSTITSCQQAPETKKPASSGSIIFGKGVKFAKDGRVTTDIIISVANEEDGRRIAAIL
NNAHYLENLHFTIDGVDTHYFVKPGPSEGDLAILGLSGRRRTLENGVNVTVSQINTML
SGTRRRYTDIQLQYRALCLNTRYGTTVDEEKRVLELARQAVRQAWRQEQRLREGE
EGLRAWTDGEKQQVLTGRVQYGDGFFVTSVEQYPELSDSANNIHFMRQSEMGR"
1813. .1881
/misc_feature
/gene="Doc4"
/note="putative; transmembrane-region site"
ORIGIN
Alignment Scores:
Pred. No.: 9722
Score: 544.00 Matches: 100
6.05e-58 Length: 9722
544.00 Matches: 100
```

```
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 1
Query Match: 99.09% Indels: 0
DB: 10 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AF059485 (1-9722)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 4069 CGGCTCTTCAGGAAGTGGTTGCCGCAGCTCCTGACCTGTCTTACTTCACTCGGAC 4128
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 4129 AAGACAGATGTCTACAACACAGAAGGTGTTTGGACTCTCAGAAGCCTTTGTTCCGTGGGT 4188
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValleuGln 60
Db 4189 TATGAGTATGATCTCTGCCAGATCTGATCCTGTGGAGAAAGGACAGCGGTGCTGCAG 4248
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 4249 GGCTATGAATGATGCCTCCAAGCTGGGGGCTGGAGTCTGGATAAGCACCATGCCCTG 4308
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 4309 AACATCCAGAGTGGCATCCTGCACAAAGGAATGGAGAGAACCAAGTTTGTGTCCACAG 4368
QY 101 Pro 101
Db 4369 CCA 4371
RESULT 4
AB025413
LOCUS
DEFINITION
AB025413
ACCESSION
AB025413.1 GI:4760781
VERSION
Ten-m4.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Ohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8585)
Ohashi,T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
Fax: +81-86-222-7768)
FEATURES
Location/Qualifiers
1..8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1..8585
/gene="ten-m4"
188..8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translation="MDVKERKPYRSLTFRRDAERRYTSSSADSEEGKGPQKSYSSSET
```

LKAYDQDARLAYGSRVKDMVPQEAEEFCRTCTNFTLRELGLGEMTPPHGTLRYRTDIGL
PHCGYSMGASSDADLEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTHEH
NTETDHPSSLQNHPRLRTPPPLPHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDH
SLSGEPPAGSAQETHAQDNWLLNSNIPLETRNLKGQPFGLTQDNLIEMDILLSASRH
DGAYSDGHFLFKPGGTSPLFCTTSPGYPLTSSVYSPPPRLPRSTFSRPAFNLLKPS
KYCNWKAAALSAILTSATLVILLAYFVAMHLFGLNWHLQPMEGQMOMYEITEDTASSW
PVPTDVSYPSSGGTLETPDRKGKGAEEKPSSLPEDSFIDSGEIDVGRRAQKIPPP
GTFWRSQVFI DHPVHLKFNVS LGKAALVGIYGRKGLPPSHTQDFVELLDGRRLLTQE
ARSLEGPPQRQSRGPVPPSSHETGFIQYLDSGIWHLAFYNDGKESEVVSFLTALIESVD
NCPNSCYNGDNCISGTCFLGFLGPDGCRASCPVLCNGQYMKGRCLCHSGWKGAE
CDVPTNQCIDVACSSHGTCIMGTCICNPGYKGESCHEVDMDPTCSSRGVCVRGECHC
SVGWGGTNCETPRATCLDQCSGHGTFELPDGLCNCDPSTGHDCSI EICAADCQGHGV
CVGGTCRCEDEGWMGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDRVV
KEGCPGLCNGRCRTLNLNGWHVCQLGWRGTGCDTSMETGCGDKGNDGDGLVDCMD
PDCCLQPLCHVNPLCLGSPDPLDIIQETQAPVSQQNLNPFYDRIKFLVGRDSTHSPG
ENPFDGGHACVIRGOVMTSDGTPLVGVNISFINNPLFGYTISRQDGSFDLVTNGGISI
ILRFERAPFITOEHTLWLPDRFFVMETIIVMRHEENELPSCDLSNFARPNPVVSPSLP
TSFASSCAEKGPVPEIQALQEEIIVAGCKMRLSYLSRTPGYKSVLRI SLTHPTIPF
NLMKVHLMVAVEGRLEFRWFAPAAADLSYFIWIDKTDVYNQKVFGEAFVSQVGEYES
CPDLILWEKRTAVLQYEIDA SKLGGWSLDKHALNIQSGILHKNGENQFVSQQPVP
IGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGDSGLYGVDFNYIRRIFFPSNV
TNILEMRNKDFRSHSPAHKYLATDPMSCGAVFLSDTNSRRVFKVSTTVVKDLVKNV
ENVAGTGQDCLPFDDTRCGDGKATEATLTNPRGITVDFKGLIYFVDGTMIRVDQNG
IISTLLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYLVLDNNVLQISE
NHQVRI VAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKIN
RIRQVTSGEISLVAGAPSGCDKNDANDCDFSGDDGYAKDAKLNTPSSLAVCAGDEL
YVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELI FDTSGKHLTYTQSLPTGDVLY
NFTYTGDDGDI THITDNNGMVNVRRDSTGMPDLWVPDGOVYVWVTGNTSALRSVTQ
GHELANMTHYHGNSGLLATKSNENGWTFYEYDSFGRLLTNVTFTQGVSSERSDTSVV
HVQVETSSKDDVTITTNLSAGAFYTLLODQVRNSYYIGADGSLRLLLLANGMEVALQT
EPHLLAGTVNPTVGRKNVTLPIDNGLNLVWRQKEQYBEI FRSLMYWMTVQYDNMGRVVKKELKVGPYAN
LDFDRVTRTEKIYDDHRKFTLRLYLDQAGRPSLWSPPSSRLNGVNVITYSPGGHIAQIOR
GIMSERMEYDQAGRITSRIFADGKMWSYTYLEKSMVLHLHSQRQYIIEFDKNDRLSSV
TMPNVARQTLTIRSVGYRNIYQPEGNASVIO DFTEDGHLHTFYLGTGRRVIYKY
GKLSKLAETLYDTTKVSFTYDETAGMLKTNLQNEGFTCTIRYQIGPLIDRQIFRFT
EEGMVNARFDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDI
NQIITTAVMTHTKHFDAYGRMKVEQYBEI FRSLMYWMTVQYDNMGRVVKKELKVGPYAN
TTRYSYEYDADGQLQTVSINDKPLWRSYDLNGLNHLSPGNSARLTPLRYDLDRIT
RLGDVQYKMEDEDFLRQDGGDVFEYNSAGLLIKAYNRASGWSVRYRYDGLGRRVSSKS
SHSHLQFFYADLTNPTKVTHLYNHSSEITSYLYDQLQHLFAMELSSGDEFYIACDN
IGTPLAVFSGTGLMIKQILYTA GEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRD
YDVLAGRWTSPPDHELWKRLSSNSIVPPHLYMFKNNNPI SNSQDIKCFMTDVSNWLTF
GFQLHNVI PGYPKPDTDAMEPSYELVHTQMKTOEWDNSKSI LGVQCEVQKQKAFVTL
ERFDQLYGSTITSQQAPETKKFASGSI FGKGVKFA LKQGRVTTDIIISVANEDGRRRI
AAILNNAHYLENLHFTIDGVDTHYFVKPSPSEGLAII LGLSGRRRTLENGVNVTSQI
NTMLSGRTRRYTDIQLQYRALCLNTRYGTTVDEEKVRLELARAQRAVQAWAREQORL
REGEEGLRAWTDGEEKQQVLNTGRVQGYDGFVTSVEQYPELSDSANNIHFMRQSEMGR
R"

ORIGIN

Alignment Scores:				
Pred. No.:	1.69e-57	Length:	8585	
Score:	540.00	Matches:	99	
Percent Similarity:	98.02%	Conservative:	0	
Best Local Similarity:	98.02%	Mismatches:	2	
Query Match:	98.36%	Indels:	0	
DB:	10	Gaps:	0	
US-10-029-020-14_COPY_1100_1200 (1-101) x AB025413 (1-8585)				
QY	1	ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp	20	
Db	3491	CGGCTCTTCAGGAAGTGGTTTGGCGCAGCTCCTGACCTGTCTATTACTTCATCTGGGAC	3550	
QY	21	LysThrAspValTyrAsnGlnLysValpPheGlyLeuSerGluAlaPheValSerValGly	40	
Db	3551	AAGACAGATGCTACAAACCAGAGAGTGTGGATTCTCAGAAGCCTTTGTTTCCGTGGGT	3610	
QY	41	TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln	60	
Db	3611	TATGAGTATGAATCCTGCCAGATCTGATCCTGTGGGAGAAAGGACAGCGGTGCTGCAG	3670	
QY	61	GlyTyrGluIleAspAlaSerLysLeuGlyTrpSerLeuAspLysHisAlaLeu	80	

Db	3671	GGCTATGAAATTGATGCCTCCAAGCTGGGGGCTGGAGTCTGGATAAGCACCATGCCCTG	3730	
QY	81	AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln	100	
Db	3731	AACATCCAGAGTGGCATCCTCGACAAAAGGAATGGAGAGAACCAAGTTTGTGTCCAGCAG	3790	
QY	101	Pro 101		
Db	3791	CCA 3793		
RESULT 5				
AX675551	AX675551	8438 bp	DNA	linear PAT 27-MAR-2003
LOCUS	Sequence 1 from Patent WO02055704.			
DEFINITION	AX675551			
ACCESSION	AX675551.1			
VERSION	AX675551.1			
KEYWORDS	GI:293333552			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1			
AUTHORS	Padigaru, M., Li, L., Zerhusen, B.D., Casman, S.J., Shenoy, S., Spytek, K.A., Zhong, M., Gangolli, E.A., Burgess, C.E., Patturajan, M., Vernet, C.A., Taylor, S., Tchernev, V.T., Miller, C.E., Guo, X., Boldog, F.L., Grosse, W.M., Alsobrook, J.P., Gerlach, V., Edingermark, S., Rothenberg, M.E., Ellerman, K., Macdougall, J., Malyankar, U., Millet, I., Peyman, J., Smithson, G., Gunther, E. and Stone, D.J.			
TITLE	Proteins, polynucleotides encoding them and methods of using the same			
JOURNAL	Patent: WO 02055704-A 1 18-JUL-2002; Curagen Corporation (US)			
FEATURES	Location/Qualifiers			
source	1..8438 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"			
ORIGIN				
Alignment Scores:				
Pred. No.:	2.57e-57	Length:	8438	
Score:	538.50	Matches:	101	
Percent Similarity:	99.02%	Conservative:	0	
Best Local Similarity:	99.02%	Mismatches:	0	
Query Match:	98.09%	Indels:	1	
DB:	6	Gaps:	1	
US-10-029-020-14_COPY_1100_1200 (1-101) x AX675551 (1-8438)				
QY	1	ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp	20	
Db	3311	CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCCAGACCTGTCTATTATTTCATTGGGAC	3370	
QY	21	LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly	40	
Db	3371	AAGACAGACGTCTACACCAGAGAGTGTTTGGGCTTTCAGAAGCCTTTGTTCCGTGGGT	3430	
QY	41	TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln	60	
Db	3431	TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGTGTGCAG	3490	
QY	61	GlyTyrGluIleAspAlaSerLysLeuGlyTrpSerLeuAspLysHisAlaLeu	80	
Db	3491	GGCTATGAAATTGACGCGTCCCAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCCCTC	3550	
QY	81	AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln	99	
Db	3551	AACATTCAAAGTGGTGGCATCCTCGACAAAAGGAATGGGGAGAACCAAGTTTGTGTCTCAG	3610	
QY	100	GlnPro 101		

gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries DNA Res. 10, 35-48 (2003)

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 5804)
Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
Direct Submission
Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7

Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.

COMMENT

FEATURES

source

1. .5804
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg00055"
/tissue_type="brain"
/dev_stage="adult"
/note="vector:modified pBC SK+"

gene

/gene="mKIAA1455"

CDS

<1. .5488
/gene="mKIAA1455"
/note="CDS is predicted by in silico analysis. Start codon is not identified."
/codon_start=2
/evidence=not_experimental
/product="mKIAA1455 protein"
/protein_id="BAC65795.1"
/db_xref="GI:28972758"

/translation="LIGNVSFHYSEYGYTITRQDMFDLVANGASLTLVFERSPE
LTQHTVWIPWNVFYVMDTLVMKKEENDIPSCDLSGFVRPSPPIVSSPLSTFFRSPE
DSPIIPETQVLHEETIIPGTDLKLSYLSSRAAGYKSVLKITMTQAVIPFNLKVLHMV
AVVGRLEQKWPASPFLAYTFIWDKTDAYNQKVYGLSEAVSVGYEYESCLDLTLWEK
RTAVLOGYELDASNMGGWTLDKHVLVDVQNGILYKNGENQFISQQPPVSSIMNGR
RRSISCSNGQADGNKLLAPVALACGIDGSLYVDNFYVRRIFPSGNVTSVLELRNK
DFRHSNPAHRYLATDPVTGDLVYSDTNRRIYRPXSLTGAKDLTKNAEVAAGTGEQ
CLPFDEARCGDGKAVEATLMSPKGMADKNGLIYFVDGTMIRKVDQNGLIISTLIGSN
DLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDDNVVLQITENRQVRIAAG
RPMHCQVPEYYPVGKHAVQTLESATAIYVSGVLIYITETDEKINRIQRTVDGE
ISLVAGIPSECDKNDANCDCYQSGDGYAKDAKINAPSSLAASPDGTLIADLGNIRI
RAVSKNKLPLNSMNFYEVASPTDQELYPIDINGHQYTVSLVTDGTYLNFYSYNDNDV
TAVTDSNGNTLIRRDENRMPVRVSPDNQVIWLTIGTNGCLKSMTAQGLELVLFYTH
GNSGLLATKSDDETGWTFDYDSEGLTNVTFPTGVVTLNHLGMDKAITVDIESSSRE
EDVSITSNLSSIDSFYTMVQDLRNSYQIYDGLSRIFYASGLDSHYQTEPHVLAGTA
NPTVAKRNMTPGNGQNLVEWFRKEQAQKQKVNFGKRLVRNGRNLLSVDFDRITKT
EKIYDDHRKFLRLIAYDTSGHPTLWLPSSKLMAVNTYSSTGQIASIQRTTSEKVDY
DSQGRIVSRVPADGKTWSYTYLEKSMVLLHLSQRQYIPEYDMWDRLSAITMPSVARHT
MQTIRSIGYRNINYPPEASNAIITDYNEEGLLQTAFLGTSRRVLFKRRQTRLSEI
LYDSTRVSFTYDETAGVLKTNLQSDGFTCTIRYQIGPLIDRIQIFRSEDGMNARF
DYSYDNSRVTSMQGVINETPLPIDLYQFDDISGKVEQFGKFGVIYYDINQIISTAVM
TYTKHFDHGRKEIQEYIFRSLMYWITIYQYDNMGRVTKREIKIGFANTTKYAYEYD
VDGQLQTVVYLNKIMWRNYDLNGLNHLNLPSSARLTPLRYDLRDRITRLGDVQYRL
DEDFGLRQRTGEITFEYSSKGLLTRYVSKSGWTVIYRDGLGRRVSSKTSLGQHLQFF
YADLTPTPTRIYTHVYNHSSSEITSLYDLOQHLFAMEISSGDEFYIASDNTGTPLAVFS
SNGMLMKQIQYATYGEIYFDSNVDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRWT
TPDIEIKWKRIKIGKDPAPFNLVYFRNNNPASKIHDVKDYITDVSWLVTFFHLHNAIPG
FPVPKFDLTPEPSYELVKSQWEDVPPIFGVQQQVARQAKAFSLSGKMAEYQVSRKAG
AEQSWLWFATVKSILIGKGMVLAVSQGRVQTNVINIANEDCIKAAVLNNAFYLENLHF
TIEGKDTHYFIKTTPTPESDLGLTLRTSGRKALENGINVTVSQSTTVNNGRTRRFADVE
MQFGALALHVRVGMVTLDEEKARILEQARQARALAWAREQQRVRDGEEGARLWTEGK
ROLLSAGKVQGYDGYVYVLSVEYQPELADSANNIQLRQSEIGKR"

ORIGIN

Alignment Scores: 7.98e-47 Length: 5804
Pred. No.: 454.00 Matches: 79
Score: 91.09% Conservative: 13
Percent Similarity:

Best Local Similarity: 78.22% Mismatches: 9
Query Match: 82.70% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AK122513 (1-5804)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 494 AGACTCTTCCAGAAGTGGTTTCTCTGCCTCGCCAACTTGGCCTACACGTTTCATCTGGGAT 553

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 554 AAGACGACGCATATAATCAGAAAGTCTACGGCTTGTACAGGCGAGTTGTCCGTCGGA 613

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGln 60
Db 614 TACGAGTAGGAGTCGTGCTTGGACCTGACTCTCTGGGAAAAGAGGACTGCCGTTTGCAG 673

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 674 GGCTATGAGTTGGATGCTTCGAACATGGCGGCTGGACGTTGGACAAGCACCATTGACTG 733

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 734 GACGTTCAGAACGGTATATACTATACAAAGGAATGGAGAAATCAGTTCATCTCTCAGCAG 793

QY 101 Pro 101
Db 794 CCT 796

RESULT 8

AF195418

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

7816 bp mRNA linear ROD 27-JAN-2000
AF195418 Mus musculus ODZ3 (Odz3) mRNA, partial cds.
AF195418 Mus musculus ODZ3 (Odz3) mRNA, partial cds.
AF195418.1 GI:6760368

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 7816)

Ben-Zur,T., Feige,E., Motro,B. and Wides,R.

The mammalian Odz gene family: Homologs of a Drosophila pair rule

gene with expression implying distinct yet overlapping

developmental roles

Dev. Biol. 117, 107-120 (2000)

2 (bases 1 to 7816)

Ben-Zur,T., Motro,B. and Wides,R.

Direct Submission

Submitted (15-OCT-1999) Faculty of Life Sciences, Bar-Ilan

University, Ramat-Gan, Israel

Location/Qualifiers

1. .7816

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

<1. .7816

/gene="Odz3"

/note="synonym: Ten-m3"

<1. .7042

/gene="Odz3"

/note="similar to the Drosophila melanogaster pair rule

odd Oz/Ten-m gene sequence"

/codon_start=2

/product="ODZ3"

/protein_id="AAF28316.1"

/db_xref="GI:6760369"

/translation="NSDTPVTNTVSLPSGDNGLGGFTHENNTIDSGELDIGRAIQE

VPPGIFWRSQFLDPQFLKFNISLQKQALIGVGRKGLPPSHTQYDFVELLDGSRLLI

AREQNLVESERAGQARSVSLHEAGFIQYLDSGIWHILAFYNDGKNPEQVSFNTIVIE

SVVECPRNCHNGECVSGTCHCFPGFLGPCSRAACPVLCSGNGQYKSCRLCFSGWK

GTECDVPTTQCIDPQCGRGICIMGSCACNSGYKGENCEEADCLDPGCSNHGVCIHGE

CHCNPGWGSNCEILKTMCADQCDSGHGTYLQESGSCTCDPNWTGPDCSNEICSVDCGS
HGVCMGSSCRCEEGWTGPACNQACHPRCAEHGTCCKGCECSQGWNGEHCTIEGCPG
LCNSNGRCTLDDQNGHWCQPQGWGAGCDVAMELTCTDSKNEGDLIDCMDDPCCLQ
SSCONQPCYCRGLPDQDIIISQSLQTPSQAAKSFYDRIIGSDSTHVLPGESPFNK
SLASVIRGQVLTAADGTPLIQVNVSLHYSEGYTITRQDGMFDLVANGGASLTILFER
SPFLTQYHTVWIPWNVYVMDTLVMKKEENDIPSCDLSGFVRPSPPIIVSSPLSTFFRS
SPEDSPIIPETQVLHEETIIPGTDLKL SylSSRAAGYSVLKITMTQAVIPFNLKMHV
LMVAVVGRLFQKFFPASPNLAYTFIWDKTDAYNQYVGLSEAVSVGYEYESCLDLTL
WEKRTAVLQGYELDASNMGGWTLDKHHVLDVQNGILYKNGENQFISQQPPVSSIMG
NGRRRSISPCSCNGQADGNKLLAPVALACGIDGSLYVGFDFNYVRRIFPSPGNVTSLEL
SSNPAHRYYLATDPTGDLVSDTNTRRIRYRPKSLTGAKDLTKNAEVVAGTGEQCLPF
DEARCGDGGKAVEATLMSPKGMAIDKNGLIYFVDGTMIRKVDQNGIISTLLGSNDLTS
ARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVLDDNNVVLQITENRQVRIAAGRPMH
CQVPGEYPPVGKHAVQTTLESATAIAVSYGLVYITETDEKKINRIRQVTTDGEISLV
AGIPSECDCKNDANDCYQSGDGYAKOAKLNAPSSLAASPDGTLTYADIAGNIRIRAVS
KNKPLLNSMNFYEVASPTDQELYIFDINGTHQYTVSLVTDGDLYNFYSNDNDVTAVT
DSNGNTLRIRDPNMPVRVVSPPNQVIWLTIGTNGCLKSMTAQGLELVLFTYHGNSSG
LLATKSDETGWTTFFDYDSEGRLTNVTFPTGVTNHLGMDMKAITVDIESSSREEDVS
ITSNLSSIDSFTYTMVQDLRNSYQIGYDGLRIFYASGLDSHYQTEPHVLAGTANPTV
AKRNMTLPGENGQNLVEWRFKEQAQGVNVPGRKLRVNGRNLSSVDFDRTTKTEKIY
DDHRKFLRLIAYDTSGHPTLWLPSSKLMANVNTYSSTQIASIQRGTTSEKVDYDSQG
RIVRSVFFADGKTWSTYILEKSMVLLHLSQRQYIFEDYMDWDRLSAITMPSVARHTMQTI
RSIGYRNINYPPEESNASIITDYNEGLLQTAFLGTSRRVLKFRQTRLSEILYDS
TRVSFTYD ETAGVLKTVNLQSDGFCITIRYQIGPLIDRQIFRFSEDGMVNARFDYSY
DNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGKGVYIYDINQIISTAVMTYTK
HFDAGRIKELIQEIFRSLMWIITIQYDNMGRVTKREIKIGPFANNTKYAYEYDVDGQ
LQTVYLNKIMRWYNDLNGNLHLNPPSSARLTPLRYDLDRDRLTRLGDVQYRLDEDG
FLRQRGTEIFEYSKGLLTRYYSKSGMWTVIYRYDGLGRVRSKTSILQHLOFFYADL
TYPTRITHVYNHSSSEITSLYYDLQHLFAMEISSGDEFYIASDNTGTPLAVFSSNGL
MLKQIQTAYGEIYFDSNVDFQVLVGFHGGLYDPLTKLIHFGERDYDILAGRWTTPDI
EIKWRIGKDPAPENLYMFRNNPNASKIHDVKDYITDVSNLVTFPGHLHNAIPGFPVP
KFDLTEPSYELVKSQWQEDVPPIFGVQQQVARQAKAFLSLGKMALEVQVSRRKAGAEQS
WLWFATVKSLLIGKVM LAVSQGRVQTNVLNIANEDCIKVAAVLNNAFYLENLHFTIEG
KDTHYFIKTTTPESDLGTLRLTSGRKALENGINVTVSQSTTVNGRTRRFADVEMQFG
ALALHVRYGWTLDEEKARILEQARQALARAWAREQQQVRVDRGEGEARLWTEGEKRQLL
SAGKVQGYDGYVLSVEQYPELADSANNIQFLRQSEIGKR"

ORIGIN

Alignment Scores:		1.1e-46	Length:	7816
Pred. No.:		454.00	Matches:	79
Score:			Conservative:	13
Percent Similarity:		91.09%	Mismatches:	9
Best Local Similarity:		78.22%	Indels:	0
Query Match:		82.70%	Gaps:	0
DB:		10		

US-10-029-020-14_COPY_1100_1200 (1-101) x AF195418 (1-7816)

QY	1	ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp	20
Db	2069	AGACTCTTCAGAGTGGTTTCTCGCTCGCCAAACTTGGCCTACACGTTTCATCTGGGAT	2128
QY	21	LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly	40
Db	2129	AAGACGGACGCATATAATCAGAAAGTCTACGGCTTGTCAGAGGACGTTGTGTCCTCGGA	2188
QY	41	TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln	60
Db	2189	TACGAGTACGAGTCGTGCTTGGACCTGACTCTCTGGGAAAAGAGGACTGCCGTTTTCGAA	2248
QY	61	GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu	80
Db	2249	GGCTATGAGTTGGATGCTTCGAACATGGCGGCTGGACGTTGGACAGCACCATGACTG	2308
QY	81	AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln	100
Db	2309	GACGTTCAGAACGGTATACTATACAAAGGAAATGGAGAAAAATCAGTTCATCTCTCAGCAG	2368
QY	101	Pro 101	
Db	2369	CCT 2371	

RESULT 9

AB025412
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

AB025412 8964 bp mRNA linear ROD 08-MAY-1999
Mus musculus mRNA for Ten-m3, complete cds.
AB025412
AB025412.1 GI:4760779
Ten-m3.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Ohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T.,
Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
proteins expressed in many tissues
J. Cell Biol. (1999) In press
2 (bases 1 to 8964)
Ohashi,T.
Direct Submission
Submitted (29-MAR-1999) Toshitaka Ohashi, Okayama University
Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
Okayama, Okayama 700-8252, Japan
(E-mail: ohashi@cc.okayama-u.ac.jp, Tel:+81-86-235-7128,
Fax:+81-86-222-7768)

FEATURES
source

1. .8964
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"
1. .8964
/gene="ten-m3"
50. .8197
/gene="ten-m3"
/codon_start=1
/product="Ten-m3"
/protein_id="BAA77398.1"
/db_xref="GI:4760780"
/translation="MDVKRRPYCSLTKSRRKERRYTNSSADNEECRVPTQKSYSS
ETLKAFDHDYSRLLYGNRVKDLVHREADYTRQGNFTLRQLGVCEASATRRGVAFCAE
MGLPHRGYSISAGSDADTENEAVMSPEHAMRLWGRVKSGRSSCLSSRSNALTITDT
EHENRSDSESEQSPNNPQPTLQPLPPSHKQHPHPSITSLNRNSLTNRNQSPAP
PAALPAELQOTTPEVOLQDSWVLGSNVPLESRHFLFKTGTGTTPLFSTATPGYTMASG
SVSPPTRLPLRNTLSRSFAFKKSKYCSWRCTALCAVGVSVLLAILLSYFTAMHLF
GLNWHLQQTENDTTFENGKVNSTVPTNTVSLPSGDNKGLGGFTHENNTIDSGELDIGR
RAIQEVPFGIFWRSQFLIDQPFQFLKFNISLQKDALIGVYGRKGLPPSHTQYDFVELLD
GSRLIAREQNLVESERAGQARSVSLHEAGFIQYLDSGIWHLAFYNDGNKPEQVSFN
TIVIESVVECPRNCHNGECVSGTCHCFGFLGPDCSRACPVLCSGNGQYSGRCLC
FSGWKGTEDVPTTQCIDPQCGGRGICIMGSCAGNSYKGCNCEADCLDPGCSNHGV
CIHGECHNPGWGSNCEILKTMCADQCDSGHGTYLQESGCTCDPNWTGPDCSNEICS
VDCSGHVCMGSCRCCEEGWTGPACNQACHPRCAEHGTCCKGCECSQGWNGEHCTI
AHYLDKIYKEGCPGLCNSNGRCTLDQNGHWCVCQPGWRGAGCDVAMETLCTDSKDN
DGLIDCMPDCCLOSSCQNPYCRGLPDQDIISQSLQTPSQAAKSFYDRISFLIGS
DSTHVLPGESPENKSLASVIRGOVLTAADGTLIGNVSVFLHYSEGYTITRQDGMFDL
VANGASLTILFERSPFLTQYHTVWIPWNVFYVMDTLVMKKEENDIPSCDLSGFVRPS
PIIVSSPLSTFFRSPEDSPIIPETQVLHEETIIPGTDLKL SylSSRAAGYSVLKIT
MTQAVIPFNLKMHVLMVAVVGRLFQKFFPASP NLAYTFIWDKTDAYNQYVGLSEAVV
SVGYEYESCLDLTLWEKRTAVLQGYELDASNMGGWTLDKHHVLDVQNGILYKNGENQ
FISQQPPVVSIMGNRRRSISCPSCNGQADGNKLLAPVALACGIDGSLYVGFDFNYV
RIFPSPGNVTSVLELRNKDFRHSNPAHRYYLATDPTGDLYVSDTNTRRIRYRPKSLT
AKDLTKNAEVVAGTGEQCLPFDEARCGDGGKAVEATLMSPKGMAIDKNGLIYFVDGTM
IRKVDQNGIISTLLGSNDLTSARPLTCDTSMHISQVRLEWPTDLAINPMDNSIYVL
NVVLQITENRQVRIAAGRPMHCQVPGEYVVGKHAVQTTLESATAIAVSYSGVLYITE
TDEKKINRIRQVTTDGEISLVAGIPSECDCKNDANDCYQSGDGYAKDNLAPSSLA
ASPDGTLTYADIAGNIRIRAVSKNKLNSMNFYEVASPTDQELYIFDINGTHQYTVSL
VTGDYLNFPYSNDNDVTAVTDSNGNTLRIRDPNMPVRVVSPPNQVIWLTIGTNGC
LKSMTAQGLELVLFTYHGNSSGLLTKSDETTGWTTFFDYDSEGRLTNVTFPTGVTN
GMDKAITVDIESSSREEDVSIITSLSSIDSFTYTMVQDLRNSYQIGYDGLRNSYQIGY
GLDSHYQTEPHVLAGTANPTVAKRNMTLPGENGQNLVEWRFKEQAQGVNVPGRKLR
VNGRNLSSVDFDRTTKTEKIYDDHRKFLRLIAYDTSGHPTLWLPSSKLMANVNTYS
GQIASIQRGTTSEKVDYDSQGRIVSRVFADGKTWSTYILEKSMVLLHLSQRQYIFEYD

gene

CDS

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3394)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submision
JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.

FEATURES
source

1..3394
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI4009028"
/tissue_type="testis"
/clone_lib="TESTI4"
/note="cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:
Pred. No.: 6,01e-47 Length: 3394
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AK125869 (1-3394)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 1514 AGACTCTTCCAAAAGTGGTTTCTGCTCCCTCACCAACTTGGCCTATACATTTTCATATGGGAT 1573
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 1574 AAAACAGATGCATATAATCAGAAAGTCTATGCTATCTGAAGCTGTTGTGTCAGTTGGA 1633
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 1634 TATGAGTATGATCGTGTGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATCTGCAG 1693
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 1694 GGCTATGAATTGGATGCGTCCCAACATGGTGGTGGACATTAGATAAACATCAGTGTCTG 1753

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 1754 GATGTACAGAACGGTATACTGTACAAGGGAAACGGGAAACCAGTTTCATCTCCACGAG 1813
QY 101 Pro 101
Db 1814 CCT 1816

RESULT 11

AB040888
LOCUS Homo sapiens mRNA for KIAA1455 protein, partial cds.
DEFINITION
ACCESSION AB040888
VERSION AB040888.1 GI:7959170
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
AUTHORS Nagase,T., Kikuno,R., Ishikawa,K., Hirose,M. and Ohara,O.

MWDRLSAITMPSVARHTMQTIRSIGYYRNIYNPPESNASIIITDYNEEGLLLQTAFLGT
SRRLFKYRQRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFICTIRYRQIGPLI
DRQIFRSEDGMVNARFDYSNFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGK
FGVIYYDINQIISTAVMTYTKHFDAGHRIKBIQYEIFRSLMYWITIQYDNMGRVTKRE
IKIGFPANTTKAYEYDVGQLOTVYLNKIMWRNYDLNGLHLLNPSSARLTPLR
YDLRDRITRLGDVQYRLDEDFLQRQRTYFVYSSKGLLTVYSSKSGWTVIYRYDGL
GRRVSSKTSLGHLQFFYADLTPTTRITHVNVHSSSEITSLYDLQHLFAMEISSGD
EFYIASDNTGTPLAFFSSNGLMLKQTYTAIGHYFDSNVDFQLVFGHGLYDPLTK
LIHFERDYDILAGRWTPDIEIKRIGKDPAPENLFRNNPNPASKIHDVKDYITDV
NSWLTVFGHLNAILPGFPVPKFDLTPESYELVKSQQWEDVPPFVGVOQVARQAKAF
LSLGMKAEVQVSRKAGAEQSWLWFATVKSLLIGKVMIAVSQGRVQTNVLNLANEDCI
KVAALNNAFYLENLHFTIEGKDTHYFIKTTTTPESDLGTLRLTSGRKALENGINVTVS
QSTTVNGRTRRFADVEMQFGALALHVRVGMTLDEKARILEOARQARALARAWAREQQ
RVRDGEGEARLWTEGEKRQLLSAGKVQGYDGYVLSVEQYPELADSANNIQLRQSEI
GKR"

ORIGIN

Alignment Scores:
Pred. No.: 1.27e-46 Length: 8964
Score: 454.00 Matches: 79
Percent Similarity: 91.09% Conservative: 13
Best Local Similarity: 78.22% Mismatches: 9
Query Match: 82.70% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AB025412 (1-8964)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3203 AGACTCTTCCAGAAGTGGTTTCTGCTCCCTCCCAAACCTTGGCCTACACGTTTCATCTGGGAT 3262
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3263 AAGACGGACGCATATAATCAGAAAGTCTACGCTTGTCCAGAGGCAGTTGTCCGTCGGA 3322
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3323 TACGAGTACGAGTCTGCTTGGACCTGACTCTCTGGGAAAAGAGGACTGCCGTTTGGCAA 3382
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3383 GGCTATGAGTTGGATGCTTCGAACATGGCGGCTGGACGTTGGACAGCACCATGACTG 3442
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3443 GACGTTTCAGAACGGTATCTACTATACAAAGGAATGGAGAAAATCAGTTTCATCTCTCAGCAG 3502

QY 101 Pro 101

Db 3503 CCT 3505

RESULT 10

AK125869
LOCUS Homo sapiens cDNA FLJ43881 fis, clone TESTI4009028, highly similar
DEFINITION to Mus musculus odd Oz/ten-m homolog 3 (Drosophila) (Odz3).
ACCESSION AK125869
VERSION AK125869.1 GI:34532128
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project

TITLE

TITLE Prediction of the coding sequences of unidentified human genes.
XVII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
JOURNAL DNA Res. 7 (2), 143-150 (2000)
MEDLINE 20277482
PUBMED 10819331
REFERENCE 2 (bases 1 to 5309)
AUTHORS Ohara,O., Nagase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)

FEATURES

source
1..5309
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fh16070"
/clone_lib="pBluescriptII SK plus"
1..5309
/gene="KIAA1455"

gene

CDS
<1..5309
/gene="KIAA1455"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1455 protein"
/protein_id="BAA95979.1"
/db_xref="GI:7959171"
/translation="QNQPYCRGLPDPQDIISQSLQSPSQAAKSFYDRISFLIGSDST
HVIPEGSPFNKSLASVIRGQVLTADGTPLIGVNSFFHYPEYGYTITRQDGMFDLVAN
GGASLTLVFERSPFLTQYHTVWIPWNVYVMDTLVMKKENDIPSCDLSGFVRNPPII
VSSPLSTFFRSSPEDSPIIPETQVLHEETTIPGTDLKL SylSSRAAGYKSVLKITMTQ
SIIPFNLKVLHVMVAVGRLFQKWFPPASPNLAYTFIWDKTDAYNOKVYGLSEAVVSVG
YEYESCLDLTLWEKRTAILQGYELDASNMGWTLDKHHVLDVQNGILYKNGENQFIS
QQPPVSSIMGNRRRSISCPSCNGQADGNKLLAPVALACGIDGSLYVGDFNVVRRIF
PSGNVTSLVLELRNKFHSSNPAHRYYLATDPVTGDLVYSDTNTRRIYRPKSLTCAKD
LTKNAEVVAGTGEQCLPEDEARCGDGGKAVEATILSPKGMADVKNGLIYFVDGTMIRK
VDQNGIISTLGSNDLTSARPLTCDTSMHI SQVRLEWPTDLAINPMDNSIYVLDNNV
LQITENRQVRIAAGRPMHCQVPGVEYPVGKHAVQTLESATAIAVSYSGVLITETDE
KKINRIQVTTDGEISLVAGIPSECDCKNDANCQYSGDGYAKDAKLSAPSSLAASP
DGTLYIADLGNIRIRAVSKNKPLLNSMNFYEVASPTDQELYIFDINGTHQYTVSLVTG
DYLNFYSYNDNDITAVTDSNGNTLRIIRDPNRPVRVSPDNQVIWLTITGNGCLKS
MTAQGLELVLFYTHGNSGLLATKSDETGTTTFDYDSEGRLTNTVTFPTGVVTNLHGDM
DKAITVDIESSSREEDVSI TNLSSIDSFTYMQDQLRNSYQIGYDGLRIIYASGLD
SHYQTEPHVLAGTANPTVAKRNMTLPGENGQNLVWEFRKEQAQGVNVPGRKLRVNG
RNLVDFDRTTKTEKIYDDHRKFLLR IAYDTSCHPTLWLPSSKLMVNVVYSSTGQI
ASIQRGTTSEKVDYDQGRIVS RVFADGKTWYSYILEKSMVLLLSQRQYIPEYDMWD
RLSAITMPSVARHTMQTIRSIGYRNINYPPESNASII TDYNEEGLLLQTAFLGTSRR
VLFKYRRQTRLSEILYDSTRVSFTYDETAGVTKVNLQSDGFCITIRYRQIGPLIDRQ
IFRSEDMVNARFDYSYDNSFRVTSMOGVINETPLPIDLYQFDDISGKVEQFGKGV
IYYDINQIISTAVMTYTKHFDAGHRIKEIQEIFRSLMWITTIQYDNMGRVTKREIKI
GPFANTTKYAYEYDVVGQLQTVYLNEKIMWRNYNDLNGNLHLLNPSNSARLTPLRIDL
RDRITRLGDVQYRLDEDFGLRQGTETFEYSKGLLTRVYKSGSGWTVIYRYDGLGRR
VSSKTSLGHLQFFFYADLTYPTRI THVYNHSSSEITSLYYDLQGHLPAMEISSGDEFY
IASDNTGTPLAVSSNGMLMKIQIYTA YGEIYFDSNIDFQLVIFPHGGLYDPLTKLIH
FGERDYDILAGRWTPDIEIKRIGKDPAPFNL YMERNNNPASKIHDVKDYITDVNSW
LVTGFEHLHNAIPGFPVPKFDLTPEPSYELVKSQDDIPPIFGVQOQVARQAKAFLSL
GKMAEVQVSRRRAGGAQSWLWFATVKS LIGKVMNAV SQGRVQTNVNLIANEDCIKVA
AVLNNAFYLENLHFTIEGKDTHYFIKTTPESDLGLTRLTSGR"

ORIGIN

Alignment Scores:
Pred. No.: 9.71e-47 Length: 5309
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AB040888 (1-5309)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 711 AGACTCTTCCAAAAGTGGTTTCTCTGCCTCACCAAACCTGGCCTATACCTTTCATATGGGAT 770

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 771 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTAGTTGGA 830

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGln 60
|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 831 TATGAGTATGAGTCGTCGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 890

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyLeuGlyTyrTrpSerLeuAspLysHisAlaLeu 80
|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 891 GGCTATGAATTGGATCGTCCAACATGGGTGGTGGACATTAGATAAACATCACGTGCTG 950

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
:::|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 951 GATGTACAGAACGGTATACTGTACAAAGGAAACGGGAAACCAGTTCATCTCCAGCAG 1010

QY 101 Pro 101
|||
Db 1011 CCT 1013

RESULT 12

AX662357

LOCUS

Sequence 39 from Patent WO02062999.

DEFINITION

AX662357

ACCESSION

AX662357.1 GI:29163218

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1
Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millet,I. and Macdougall,J.R.
Proteins and nucleic acids encoding same
Patent: WO 02062999-A 39 15-AUG-2002;
Curagen Corporation (US)

TITLE

JOURNAL

Curagen Corporation (US)

FEATURES

source

1..8473
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.6e-46 Length: 8473
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AX662357 (1-8473)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 3150 AGACTCTTCCAAAAGTGGTTTCTCTGCCTCACCAAACCTGGCCTATACCTTTCATATGGGAT 3209

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 3210 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTAGTTGGA 3269

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGln 60
|||||:::||||| ||:::||||| ||:::||||| ||:::||||| ||:::|||||
Db 3270 TATGAGTATGAGTCGTCGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 3329

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisHisAlaLeu 80
Db 3330 GGCTATGAATTGGATCGCTCCAAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 3389
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3390 GATGTACAGAACGGTATACTGTACAAGGAAACGGGGAAAAACCAAGTTTCATCTCCAGCAG 3449
QY 101 Pro 101
Db 3450 CCT 3452
RESULT 13
AX662359
LOCUS AX662359 8487 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 41 from Patent WO02062999.
ACCESSION AX662359
VERSION AX662359.1 GI:29163219
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millet,I. and Macdougall,J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 41 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1. .8487
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 1.6e-46 Length: 8487
Pred. No.: 453.00 Matches: 78
Score: 91.09% Conservative: 14
Percent Similarity: 77.23% Mismatches: 9
Best Local Similarity: 82.51% Indels: 0
Query Match: 6 Gaps: 0
DB: 6
US-10-029-020-14_COPY_1100_1200 (1-101) x AX662359 (1-8487)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3146 AGACTCTTCCAAAAGTGGTTTCCTGCCTCACCAAACTTGGCCTATACTTTTCATATGGGAT 3205
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3206 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTCAGTTGGA 3265
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3266 TATGAGTATGATCGTGTGTTGGACCTGACTCTGTGGGAAAAAGAGGACTGCCATTCTGCAG 3325
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisHisAlaLeu 80
Db 3326 GGCTATGAATTGGATCGCTCCAAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 3385
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3386 GATGTACAGAACGGTATACTGTACAAGGGAAACGGGGAAAAACCAAGTTTCATCTCCAGCAG 3445
QY 101 Pro 101
Db 3450 CCT 3452

Db 3446 CCT 3448
RESULT 14
AX662355
LOCUS AX662355 8645 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 37 from Patent WO02062999.
ACCESSION AX662355
VERSION AX662355.1 GI:29163217
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G.,
Millet,I. and Macdougall,J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 37 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1. .8645
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 1.64e-46 Length: 8645
Pred. No.: 453.00 Matches: 78
Score: 91.09% Conservative: 14
Percent Similarity: 77.23% Mismatches: 9
Best Local Similarity: 82.51% Indels: 0
Query Match: 6 Gaps: 0
DB: 6
US-10-029-020-14_COPY_1100_1200 (1-101) x AX662355 (1-8645)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3322 AGACTCTTCCAAAAGTGGTTTCCTGCCTCACCAAACTTGGCCTATACTTTTCATATGGGAT 3381
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3382 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTCAGTTGGA 3441
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3442 TATGAGTATGATCGTGTGTTGGACCTGACTCTGTGGGAAAAAGAGGACTGCCATTCTGCAG 3501
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisHisAlaLeu 80
Db 3502 GGCTATGAATTGGATCGCTCCAAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 3561
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3562 GATGTACAGAACGGTATACTGTACAAGGAAACGGGGAAAAACCAAGTTTCATCTCCAGCAG 3621
QY 101 Pro 101
Db 3622 CCT 3624
RESULT 15
AX662353
LOCUS AX662353 8675 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 35 from Patent WO02062999.
ACCESSION AX662353
VERSION AX662353.1 GI:29163216
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X., Malyankar,U., Shimkets,R.A., Tchernev,V.T., Spaderna,S.K., Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A., Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F., Burgess,C.E., Edinger,S., Ellerman,K., Gunther,E., Smithson,G., Millet,I. and Macdougall,J.R.

Proteins and nucleic acids encoding same Patent: WO 02062999-A 35 15-AUG-2002; Curagen Corporation (US)

FEATURES

source 1..8675 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1.64e-46 Length: 8675 Score: 453.00 Matches: 78 Percent Similarity: 91.09% Conservative: 14 Best Local Similarity: 77.23% Mismatches: 9 Query Match: 82.51% Indels: 0 DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AX662353 (1-8675)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20 Db 3334 AGACTCTTCCAAAAGTGGTTTCCTGCCCTCACCAAACTTGGCCTATACCTTCATATGGGAT 3393 QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40 Db 3394 AAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTGTGTCAGTTGGA 3453 QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60 Db 3454 TATGAGTATGAGTCGTCGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 3513 QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80 Db 3514 GGCTATGAATTGGATGCGTCCACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 3573 QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100 Db 3574 GATGTACAGAACGGTATACTGTACAGGGAAACGGGGAAACAGTTTCATCTCCACGACG 3633 QY 101 Pro 101 Db 3634 CCT 3636

Search completed: August 14, 2004, 11:52:10 Job time : 1809.37 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 9568.56 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
Perfect score: 2845
Sequence: 1 YYIGADGSLRLLLANGMEVA.....YNRAGSWSVRYRDGLGRRV 541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2 1/USPTO spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN 1 1 19065 @runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Result No.	Score	Query Match %	Length	DB ID	Description
1	2845	100.0	3048	9	AK127705 Homo sapi
2	2845	100.0	8354	6	AX556500 Sequence
3	2845	100.0	8624	9	AB037723 Homo sapi
4	2845	100.0	8645	6	AX600210 Sequence
5	2845	100.0	8993	9	BX640737 Homo sapi
6	2840	99.8	3320	9	AK127101 Homo sapi
7	2834	99.6	3486	9	BX537983 Homo sapi
8	2821.5	99.2	8438	6	AX675551 Sequence
9	2813	98.9	5583	10	AK122490 Mus muscu
10	2813	98.9	8585	10	AB025413 Mus muscu
11	2785	97.9	9722	10	AF059485 Mus muscu
12	2573	90.4	9264	5	AB026980 Danio rer
13	2233	78.5	3614	6	AX876525 Sequence
14	2233	78.5	3614	6	BD156175 Primer fo
15	2233	78.5	3614	9	AK001336 Homo sapi
16	2233	78.5	5309	9	AB040888 Homo sapi
17	2233	78.5	8473	6	AX662357 Sequence
18	2233	78.5	8487	6	AX662359 Sequence
19	2233	78.5	8645	6	AX662355 Sequence
20	2233	78.5	8675	6	AX662353 Sequence
21	2228	78.3	5804	10	AK122513 Mus muscu
22	2228	78.3	7816	10	AF195418 Mus muscu
23	2228	78.3	8964	10	AB025412 Mus muscu
24	2219	78.0	170156	2	AP002412 Homo sapi
25	2199.5	77.3	186084	9	AP002768 Homo sapi
26	2188	76.9	8816	5	AB026979 Danio rer
27	2180.5	76.6	208760	2	AC074048 Mus muscu
28	2180.5	76.6	226256	10	AC073599 Mus muscu
29	2165	76.1	231970	2	AC133064 Rattus no
30	2165	76.1	257693	2	AC103323 Rattus no
31	2159	75.9	3270	6	AX877449 Sequence
32	2159	75.9	3270	6	BD156663 Primer fo
33	2159	75.9	3270	9	AK001748 Homo sapi
34	2141.5	75.3	8797	6	AX250063 Sequence
35	2141.5	75.3	8797	6	AX250066 Sequence
36	2141.5	75.3	8797	10	AB025411 Mus muscu
37	2138.5	75.2	8409	5	GGA279031 Gallus ga
38	2138.5	75.2	8409	6	AX250068 Sequence
39	2136.5	75.1	6560	6	AX250065 Sequence
40	2136.5	75.1	7781	9	AB032953 Homo sapi
41	2136.5	75.1	9729	6	AX250013 Sequence
42	2136.5	75.1	9826	6	AX250008 Sequence
43	2135.5	75.1	8689	6	AX250067 Sequence
44	2135.5	75.1	8689	10	AF086607 Rattus no
45	2124.5	74.7	7713	9	BX648178 Homo sapi

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

AK127705 AK127705 3048 bp mRNA linear PRI 09-SEP-2003
LOCUS Homo sapiens cDNA FLJ45805 fis, clone NT2RI3007095, highly similar
DEFINITION to Mus musculus neuregulin 1 (Nrg1).
ACCESSION AK127705
VERSION AK127705.1 GI:34534730
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S.,
AUTHORS Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A.,
Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K.
and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3048)
REFERENCE Isogai,T. and Yamamoto,J.
AUTHORS Direct Submission
TITLE Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
JOURNAL Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source Location/Qualifiers
1..3048
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RI3007095"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RI3"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells treated 2-weeks mitotic inhibitor after
5-weeks retinoic acid (RA) induction.-majorly NT2 neuron"
ORIGIN
Alignment Scores:
Pred. No.: 1.01e-236 Length: 3048
Score: 2845.00 Matches: 541
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x AK127705 (1-3048)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db 670 TACTACATCGGGCCGATGGCTCCTTCGGCTGCTGCTGCCAACGGCATGGAGGTGGCG 729
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 730 CTGCAGACTGAGCCCCACITGTGGCTGGCACCGTCAACCCACCGTGGGCAAGAGGAAT 789
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 790 GTCAGCTGCCCATCGACACGGCCTCAACCTGGTGGAGTGGGCCAGCGCAAGAGCAG 849
QY 61 AlaArgGlyGlnValThrValPheArgLeuArgLeuValHisAsnArgAsnLeuLeu 80

Db 850 GCTCGGGGCCAGGTCACTGTCTTTGGGCGCCGGCTGCGGTTTCAAAACCGAAATCTCCTA 909
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 910 TCTCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATGATGACCAACCGCAAGTTC 969
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 970 ACCCTTCGGATTCTGTACGACCAGCGGGGGCGCCAGCCCTCTGGTCAACCCAGCAGCAG 1029
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 1030 CTGAATGGTGTCAACGTGACATACTCCCTGGGGTTACATTGCTGGCATCCAGAGGGGC 1089
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 1090 ATCATGTCTGAAAGAATGGAATACGACCAGCGGGCGCCGCATCACATCCAGGATCTTCGCT 1149
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 1150 GATGGGAAGACATGGAGCTACACATACTAGAGAAGTCCATGGTGTGCTACTACACAGC 1209
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 1210 CAGAGGCAGTATATCTTTGAGTTCGCAAGAATGACCGCCTCTCTTCTGTGACGATGCCC 1269
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 1270 AACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT 1329
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 1330 CAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGGCACCTCCTT 1389
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 1390 CACACCTTCTACTGGGCACCTGGCCGAGGTGATATACAAGTATGGCAAACTGTCAAAG 1449
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 1450 CTGGCAGAGACGCTCTATGACACCACCAAGGTCAGTTTACCTATGACGAGACGGCAGGC 1509
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 1510 ATGCTGAAGACCCTCAACCTACAGAAATGAGGGCTTCACCTGCACCATCCCGTACCGTCAG 1569
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 1570 ATTGGGCCCCCTGATTGACCGCAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCC 1629
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 1630 CGTTTTGACTACAACCTATGACAACAGCTTCCGGGTGACCCAGCATGCAGGCTGTGATCAAC 1689
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db 1690 GAGACCCCACTGCCCAFTGATCTCTATCGCTATGATGATGTGTCTCAGGCAAGACAGAGCAG 1749
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 1750 TTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACCATCATCACACAGCTGTTCATG 1809
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 1810 ACCCACCAAGCATTTTGATGCATATGGCAGGATGAGGAAGTGCAGTATGAGATCTTC 1869
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 1870 CGTCGCTCATGTACTGGTACCGTCCAGTATGATGATGATGAGGGCGAGTAGTGAAGAAG 1929
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440


```
Db      1930  GAGCTGAAGGTAGGACCCCTACGCCAATACCACCTCGCTACTCCTATGAGTATGATGCTGAC 1989
Qy      441  GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db      1990  GGCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC 2049
Qy      461  AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db      2050  AATGGAAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTAT 2109
Qy      481  AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db      2110  GACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGC 2169
Qy      501  PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db      2170  TTCCTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTGCTCATCAAG 2229
Qy      521  AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db      2230  GCCTACAACCGGGCTGGCAGCTGGAGTGTCAAGTACCGTACGATGGCCTGGGGCGCGC 2289
Qy      541  Val 541
Db      2290  GTG 2292

RESULT 2
AX556500
LOCUS      AX556500                      8354 bp      DNA      linear      PAT 27-NOV-2002
DEFINITION Sequence 13 from Patent WO02057453.
ACCESSION  AX556500
VERSION     AX556500.1  GI:25899736
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS    Gangolli, E.A., Patturajan, M., Vernet, C.A., Malyankar, U.M.,
            Kekuda, R., Stone, D.J., Anderson, D., Shimkets, R.A., Burgess, C.E.,
            Zerhusen, B.D., Liu, X., Spytek, K.A., Casman, S.J., Boldog, F.L.,
            Smithson, G., Li, L. and Ji, W.
TITLE      Polypeptides and nucleic acids encoding same
JOURNAL    Patent: WO 02057453-A 13 25-JUL-2002;
            Curagen Corporation (US)
FEATURES
            Location/Qualifiers
            source          1..8354
                           /organism="Homo sapiens"
                           /mol_type="unassigned DNA"
                           /db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.:      3.81e-236      Length:      8354
Score:          2845.00      Matches:      541
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             6      Gaps:      0

US-10-029-020-14_COPY_1760_2300 (1-541) x AX556500 (1-8354)

Qy      1  TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db      5312  TACTACATCGGGCCGATGGCTCCCTCGGGCTGCTGCTGGCCACGGCATGGAGGTGGCG 5371
Qy      21  LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db      5372  CTGCAGACTGAGCCCCACTTGTCTGGCTGGCACCGTCAACCCACCGTGGGCAAGAGGAAT 5431
Qy      41  ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db      5432  GTACAGCTGCCCCATCGACAACCGGCCTCAACCTGGTGGAGTGGGCGCCAGCGCAAGAGCAG 5491
```

```
Qy      61  AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db      5492  GCTCGGGGCCAGGTCACTGCTTTTGGGCGCCGGCTGCGGGTGCACAACCGAAATCTCCTA 5551
Qy      81  SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db      5552  TCTCTGGACTTTGATCGCGTAAACAGCACAGAGAAGATCTATGATGACCAACCGCAAGTTC 5611
Qy      101  ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db      5612  ACCCTTCGGATTCTGTACGACGCGGGCGGCCAGCTCTGGTCACCCAGCAGCAGG 5671
Qy      121  LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db      5672  CTGAATGGTGTCAACGTGACATACTCCCTTGGGGTTACATTGCTGGCATCCAGAGGGGC 5731
Qy      141  IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db      5732  ATCATGTCTGAAGAATGGAATACGACCAGCGGGCGGCATCACATCCAGGATCTTCGCT 5791
Qy      161  AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db      5792  GATGGGAAGACATGGAGCTACACATACTTAGAGAAGTCCATGGTGTCTACTACACAGC 5851
Qy      181  GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db      5852  CAGAGGCAGTATATCTTTGAGTTCGACAAGAATGACCGCCTCTCTTCTGTGACGATGCCC 5911
Qy      201  AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db      5912  AACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT 5971
Qy      221  GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db      5972  CAGCCCCCTGAGGCAATGCCTCAGTCAATACAGGACTTCACCTGAGGATGGCACCTCCTT 6031
Qy      241  HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db      6032  CACACCTTCTACCTGGGCACCTGGCCGCGCGGTGATATACAAGTATGGCAAACTGTCAAAG 6091
Qy      261  LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db      6092  CTGGCAGAGACGCTCTATGACACCACCAAGGTGAGTTTCACCTATGACGAGACGGCAGGC 6151
Qy      281  MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db      6152  ATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACCTGCACCATCCGCTACCGTCAG 6211
Qy      301  IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db      6212  ATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCACCTGAGGAAGGCATGGTCAACGCC 6271
Qy      321  ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db      6272  CGTTTTGACTACAACCTATGACAACAGCTTCCGGGTGACCAGCATGCAGGCTGTGATCAAC 6331
Qy      341  GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db      6332  GAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGTGAGGCAAGACAGAGCAG 6391
Qy      361  PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db      6392  TTTGGGAAGTTTGGTGTCACTTACTATGACATTAACCAGATCATCACACAGCTGTTCATG 6451
Qy      381  ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db      6452  ACCCACCAAGCATTTTGTGTCATATGAGGATGAAGGAAGTGCAGTATGAGATCTTC 6511
Qy      401  ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db      6512  CGTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAGAG 6571
```

421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
6572	GAGCTGAAGGTAGGACCCCTACGCCAATACCACCTCGCTACTCTATGAGTATGATGCTGAC	6631
441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu	460
6632	GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC	6691
461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
6692	AATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTAT	6751
481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
6752	GACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAGATGGATGAGGATGGC	6811
501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
6812	TTCTTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCGTCTCATCAAG	6871
521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
6872	GCCTACAACCGGGCTGGCAGCTGGAGTGTGAGGTACCGTACCGTACGATGGCTGGGGCGGCGC	6931
541	Val	541
6932	GTG	6934
RESULT 3		
8037723		
OCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
MEDLINE		
PUBMED		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
source		
gene		
CDS		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		

Db 5901 CAGAGGCAGTATATCTTTGAGTTCGACAGAATGACCGCTCTCTTCTGTGACGATGCC 5960

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 5961 AACGTGGCGGCAGACACTAGACACCATCCGCTCAGTGGCTACTACAGAAACATCTAT 6020

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 6021 CAGCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACCTGAGGATGGCACCTCCTT 6080

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrIlystTyrGlyLysLeuSerLys 260

Db 6081 CACACCTTCTACCTGGGCACCTGGCCGCGAGGTGATATACAAGTATGGCAAACTGTCAAAG 6140

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 6141 CTGGCAGAGACGCTCTATGACACCACCAAGGTCAGTTTACCTATGACGAGACGGCAGGC 6200

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db 6201 ATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACCTGCACCATCCGCTACCGTTCAG 6260

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320

Db 6261 ATTGGGCCCCGATTGACCGACAGATCTTCGCTTCACTGAGGAAGGCATGGTCAACGCC 6320

QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340

Db 6321 CGTTTTGACTACAACTATGACAAACAGCTTCCGGGTGACCAAGCATGCAGGCTGTGATCAAC 6380

QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360

Db 6381 GAGACCCCACTGCCCATGATCTCTATCGCTATGATGATGTGTGAGGCAACAGACAGCAG 6440

QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380

Db 6441 TTGGGAAGTTTGGTGTCAATTACTATGACATTAAACAGATCATCACACAGTGTGATG 6500

QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400

Db 6501 ACCCACCAAGCATTTTGTATGTCATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTC 6560

QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420

Db 6561 CGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGCGGAGTAGTGAAGAAG 6620

QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440

Db 6621 GAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCCTATGAGTATGATGCTGAC 6680

QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460

Db 6681 GGCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC 6740

QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480

Db 6741 AATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCAAGGCTCACACCACCTACGGTAT 6800

QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500

Db 6801 GACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGC 6860

QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520

Db 6861 TTCCTGAGCAGCGGGCGGTGATATCTTGAGTACAACCTACGCTACGCTGGCCTGCTCATCAAG 6920

QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540

Db 6921 GCCTACACCGGGCTGGCAGCTGGAGTGTTCAGGTACCGCTACGATGGCCTGGGGCGGCGC 6980

QY 541 Val 541

Db 6981 GTG 6983

RESULT 5

HSM806812 8993 bp mRNA linear PRI 28-AUG-2003

LOCUS Homo sapiens mRNA; cDNA DKFZp686K11107 (from clone DKFZp686K11107).

DEFINITION BX640737

ACCESSION BX640737.1 GI:34364828

VERSION

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 8993)

AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.

CONSRMT The German Human cDNA Consortium

TITLE Direct Submission

JOURNAL Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp686K11107) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.

Location/Qualifiers

1..8993

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/map="11q14.1"

/clone="DKFZp686K11107"

/tissue_type="human fetal kidney"

/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfilA + SfilB"

/dev_stage="fetal"

1..8993

/gene="DKFZp686K11107"

<1..4200

/gene="DKFZp686K11107"

/note="hypothetical protein, N-terminus elongated, differentially spliced"

/codon_start=1

/product="hypothetical protein"

/protein_id="CAE45850.1"

/db_xref="GI:34364829"

/translation="RIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVHLEWPTDLAI NPMDNSLYLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESAlA LAVSHNGVLYIAETDEKKINRIQVTTSGEISLVAGAPSGCDCKNDANCDFSGDDGY AKDAKLNTPSSLAVCADGELYVADLGNIRIRIRKKNKPFNTQNMVELSSPIDQELYL FDTTGKHLTYQSLPTGDYLYNFTYTGDDITLITDNNGMVNVRRDSTGMPLWLVPD GOVYVVTMTGNSALKSVTTQGHELAMMTYHGSGLLATKSNENGWTTFFYEYDSFGRLT NVTFTGQVSSFRSDTSSVHVQVETSSKDDVTITNLNSASGAFYTLQDQVNSYYI GADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGNKRVNTPIDNGLNLVEWRQKEQA RGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIDYDHRKFTLRILYDQAGRPSLWSPSS RLNGVNTYSPGGYIAGIQGIMSERMEYDQAGRITSRIADGKTWSYTYLEKSMVLL LHSQRYIFEFDKNDRSLSSVTMPNVARQTLTIRSVGYVRNIYQPEGNASVTDQFTE DGHLLHTFYLGTRRVIYKYGKLSKLAETLYDTTKTFTYDETAGMLKTINLQNEGFT CTIRYQIGPLIDRQIFRFTTEEGMVARFDYNDNSFRVTSMQAVINETPLPIDLYRY DDVSGKTEQFGKFGYIYDINQIITTAVMTHTFDAYGRMKVEVQYEIFRSLMYMTIV QYDNMGRVVKKELKVGPYANTTRYSEYEDADGQLQTVSINDKPLWRSYDNLGNLHL SPGNSARLTPLRYDIRITRLGDVQYKMDDEGFLRQRGGDIFEYNSAGLLIKAYNRA GSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPYKVTLYNHSSEITSLYYDLQ GHLFAMELSSGDEFYIACDNIGTPLAVSGTGLMKIQLIYAYGEIYMTDPNFQIIL GYHGLGLYDPLTKLVHMRDRDYDVLAGRWTSPDHFWKHLSSSNVMPFNLYMFKNNPI SNSQDIKCFMTDVNSWLLTFGFLHNHVI PGYPKPDMDAMEPSYELIHTQMKTQEWDNS KSLILGVQCEVQKQLKAFVTILRFDQLYGSTITSCQQAQPKTKKFASSGGSVFGKGVKFAI

KDGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPSPSEGLAIL
GLSGGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDEKARV
LELARQRAVRQAWAREQOIRLEEGEGLRAWTEGEKQVILSTGRVQYDGFVVISVEQY
PELSDSANNIHFMRSQSEMGRR"
polyA_signal 8953. .8958
/gene="DKFZp686K11107"
polyA_site 8973
/gene="DKFZp686K11107"

ORIGIN

Alignment Scores:

Pred. No.: 4.19e-236 Length: 8993
Score: 2845.00 Matches: 541
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x HSM806812 (1-8993)

QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla	20
Db	1168	TACTACATCGGGCGCGATGGCTCCTTGC	1227
QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn	40
Db	1228	CTGCAGACTGAGCCCCACTTGCTGGCTGGCACCCGTCACACCCACCGTGGGCAAGAGGAAT	1287
QY	41	ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln	60
Db	1288	GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCCAGCGCAAGAGCAG	1347
QY	61	AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu	80
Db	1348	GCTCGGGCCAGGTCACTGTCTTTGGCGCCCGGCTGGGGTTTCAACCGAAATCTCCTA	1407
QY	81	SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe	100
Db	1408	TCTCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATGATACCCACCGCAAGTTC	1467
QY	101	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg	120
Db	1468	ACCCCTTCGGATTCTGTACGACCAGCGGGCGGCCAGCCTCTGGTCACCAGCAGCAGG	1527
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly	140
Db	1528	CTGAATGGTGTCAACGTGACATACTCCCTGGGGGTTACATTGCTGGCATCCAGAGGGC	1587
QY	141	IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla	160
Db	1588	ATCATGTCTGAAGAATGGAATACGACAGCGCGGCCGCATCACATCCAGGATCTTCGCT	1647
QY	161	AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer	180
Db	1648	GATGGGAAGACATGGAGCTACACATACTTAGAGAAATGACCGCCTCTCTTGTGACGATGCC	1707
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro	200
Db	1708	CAGAGCAGTATATCTTTGAGTTCGACAAGAATGACCGCCTCTCTTGTGACGATGCC	1767
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr	220
Db	1768	AACGTGGCGCGGCAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTAT	1827
QY	221	GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu	240
Db	1828	CAGCCCTCTGAGGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGCACCTCCTT	1887
QY	241	HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys	260
Db	1888	CACACCTTCTACCTGGGCACTGGCGCAGGGTGATATACAAGTATGGCAAACTGTCAAG	1947
QY	261	LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly	280

Db	1948	CTGGCAGACGCTCTATGACACCAAGGTCAGTTTCACCTATGACGAGACGGCAGC	2007
QY	281	MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln	300
Db	2008	ATGCTGAAGACCATCAACCTACAGAATGAGGGCTTCACCTGCACCATCCGCTACGTCAG	2067
QY	301	IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla	320
Db	2068	ATTGGGCCCTTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCC	2127
QY	321	ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn	340
Db	2128	CGTTTGACTACAACATATGACAACAGCTTCCGGGTGACCAGCATGCAGGCTGTGATCAAC	2187
QY	341	GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln	360
Db	2188	GAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGAGCAAGACAGAGCAG	2247
QY	361	PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet	380
Db	2248	TTTGGGAAGTTTGGTGTCAATTTACTATGACATTAAACAGATCATCAACACAGCTGTCTATG	2307
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
Db	2308	ACCCACACCAAGCAFTTTGATGCATATGGCAGATGAAGGAAGTGCAGTATGAGATCTTC	2367
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420
Db	2368	CGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAAATGAGGCGAGTAGTGAAGAAG	2427
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
Db	2428	GAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGATGTGAC	2487
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu	460
Db	2488	GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC	2547
QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
Db	2548	AATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTAT	2607
QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
Db	2608	GACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGC	2667
QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuIleLys	520
Db	2668	TTCCCTGAGCGAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCCTGCTCATCAAG	2727
QY	521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
Db	2728	GCCTACAACCGGGCTGGCAGCTGGAGTGTCAGGTACCGCTACGATGGCCTGGGGCGCGC	2787
QY	541	Val 541	
Db	2788	GTG 2790	

RESULT 6

AK127101

LOCUS

DEFINITION

AK127101

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

AK127101 3320 bp mRNA linear PRI 09-SEP-2003
Homo sapiens cDNA FLJ45158 fis, clone BRAWH3043034, highly similar
to Mus musculus neuregulin 1 (Nrg1).

AK127101 GI:34533862
oligo capping; fis (full insert sequence).

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,

Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J.,
Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K.,
Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3320)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .3320
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAWH3043034"
/tissue type="brain"
/clone_lib="BRAWH3"
/note="cloning vector: pME18SFL3"

Alignment Scores:
Pred. No.: 3 06e-236 Length: 3320
Score: 2840.00 Matches: 540
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 99.82% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AK127101 (1-3320)

1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
36 TACTACATCGGGCCGATGGCTCCTTGGGCTGCTGGTGGCCAAACGGCATGGAGTGGCG 95
21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
96 CTGACAGCTGAGCCCCACTTGTGGCTGGCACCGTCAACCCACCGTGGGCAAGAGGAAT 155
41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgLysGluGln 60
156 GTCAGCTGCCCATCGACACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAG 215
61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
216 GTCGGGGCCAGGTCACTGTCCTTGGGGCCGGCTGGGGGTTCACAACCGAAATCTCCTA 275
81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
276 TCTCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATGATGACCACCGCAAGTTC 335
101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
336 ACCCTTCGGATTCTGTACGACCAAGCGGGCGGCCAGCCCTCTGTCACCCAGCAGCAGG 395
121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
396 CTGAATGGTGTCAACGTGACATACTCCCTGGGGGTTACATTGCTGGCATCCAGAGGGGC 455
141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

456 ATCATGTCTGAAAGAATGGAATACGACCAGCGGGCCGCATCACATCCAGGATCTTCGCT 515
161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
516 GATGGGAAGACATGGAGCTACACATACTTAGAGAAGTCCATGGTGTCTACTACACAGC 575
181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
576 CAGAGGCAGTATATCTTTGAGTTCGACAAGAATGACCGCCTCTCTGTGACGATGCC 635
201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
636 AACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT 695
221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
696 CAGCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGCACCTCCT 755
241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
756 CACACCTTCTACCTGGGCACCTGGCGCAGGGTGATATACAAGTATGCAAACTGTCAAAG 815
261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
816 CTGGCAGAGACGCTCTATGACACCCCAAGGTCAGTTTCACCTATGACGAGACGGCAGGC 875
281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
876 ATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACCTGCACCATCCGCTACCGTCAG 935
301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
936 ATTGGGCCCCGTGATTGACCGCAGACATCTTCGTTTCACTGAGGAAGGCATGGTCAACGCC 995
321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
996 CGTTTGTGACTACAACTATGACAACAGCTTCCGGGTGACCAGCATGCAGGCTGTGATCAAC 1055
341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
1056 GAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGTGTGAGGCAAGACAGAGCAG 1115
361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
1116 TTTGGGAAGTTTGGTGTCAATTTACTATGACATTAAACCAGATCATCACACAGCTGTCTATG 1175
381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
1176 ACCCACACCAAGCATTTTGATGCATATGGCAGGATGAAGAAAGTGAGTATGAGATCTTC 1235
401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
1236 CGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAAACATGGGGCGAGTAGTGAAGAAG 1295
421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
1296 GAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGATGCTGAC 1355
441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
1356 GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC 1415
461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
1416 AATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTAT 1475
481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
1476 GACATCCGCGACCGCATCACTCGCTGGGTGACCGTGCATATACAAGATGATGAGGATGGC 1535
501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
1536 TTCCTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTGCTCATCAAG 1595

Qy 521 AlatyAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
|||||
Db 1596 GCCTACAAACCGGGCTGGCAGCTGGAGTGTACAGTACCGCTACGATGGCCTGGGGCGGCGC 1655
Qy 541 Val 541
|||
Db 1656 GTG 1658

RESULT 7
HSM806114
LOCUS
DEFINITION Homo sapiens mRNA; cDNA DKFZp686D0412 (from clone DKFZp686D0412).
ACCESSION BX537983
VERSION BX537983.1 GI:31874053

SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3486)

AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
Direct Submission
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.

TITLE This clone (DKFZp686D0412) is available at the RZPD in Berlin.
JOURNAL Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
Source Location/Qualifiers
1..3486
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="llql4.1"
/clone="DKFZp686D0412"
/tissue type="human cervix"
/clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host
DH10B; sites SfiIA + SfiIB"
/dev stage="adult"
1..3486
/gene="DKFZp686D0412"
<1..3217
/gene="DKFZp686D0412"
/note="hypothetical protein, N-terminus truncated"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAD97943.1"
/db_xref="GI:31874054"
/translation="SFGRLTNVTFPTGQVSSFRSDTSSVHVQVETSSKDDVTITNL
SASGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRNV
TLPIDNGLNLVEWRQKEARGQVTVFGRRLRVNENLLSLGFDVRVTRTEKIYDDHRK
FTLRILYDQAGRPSLWSPSSRLNGVNVITYSPGGYIAGIQGIMSERMEYDQAGRITSR
IFADGKTWSYTYLEKSMVLLHSQRQYIFEEFKNDRLLSSVTMPNVAQTLETIRSVGY
YRNIYQPPEGNASVIQDFTEDGHLHTFYLGTRGVYKYGKLSKLAETFDYDTTKVSF
TYDETAGMLKTNLQNEGFTCTIRYQIGPLIDRQIFRTEEGMVNARFDYNDNSFR
VTSMQAVINETPLPIDLYRYDDVSGKTEQFGKGVIIYDINQIITTAVMTHTKHFDAY
GRMKVQVEIFRSLMYMTVQYDNMGRVVKKELKVGYPYANTTRYSEYVDADGLOQTVS
INDKPLWRYSDYDNLGNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDGDFLRQR
GGDIFEYNSAGLLIKAYNRAGSWSVRYRVDGLGRRVSSKSSHLLQFFYADLTNPTK
VTHLYNHSSESSEITSLYDYLQHLFAMELSSGDEFYIACDNIPTPLVFSGTGLMIKQI
LYTAYGIEYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDLAGRWTSPDHELWKH
LSSSNVMPFNLVYFKNPNPISNSQDIKCFMTDVNSWLLTFGFLHNVIPGYPKPDMDA
MEPSYELIHTQMTQEWDNKSILGVQCEVQKQKAFVTLEFRFDQLYGSTITSCQAP
KTKKFASSGSVFGKGVKFAKXGRVTTDIISVANEDGRRVAAAILNHAHYLENLHFTID

GVDTHYFVKPGPSEGLAILGLSGGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQY
GALCLNTRYGTTLDEKARVLELARQAVRQAWAREQRLREGEGLRAWTEGEKQOV
LSTGRVQGYDGFVVISVEQYPELSDSANNIHFMQRQSEMGRR"
3468
/gene="DKFZp686D0412"

polyA_site
ORIGIN
Alignment Scores:
Pred. No.: 1.08e-235 Length: 3486
Score: 2834.00 Matches: 539
Percent Similarity: 99.63% Conservative: 0
Best Local Similarity: 99.63% Mismatches: 2
Query Match: 99.61% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x HSM806114 (1-3486)
Qy 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db 185 TACTACATCGGGCGCGATGGCTCCTTGGCGTGTCTGGTGGCCACGGCATGGAGGTGGCG 244
Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 245 CTGCAGACTGAGCCCCACTTGTGGCTGGCACCGCTCAACCCACCCTGGGCAAGAGGAAT 304
Qy 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 305 GTCACGCTGCCCATCGACAACGGCTCAACCTGGTGAGTGGCGCCAGCGCAAAAGAGCAG 364
Qy 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 365 GCTCGGGGCCAGGTCACTGTCTTTGGCGCGCGGTGCGGGTTTCACAAACGAAATCTCCTA 424
Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 425 TCTCTGGGCTTTGATCGCGTAACACGCACACAGAGAAGATCTATGATGACCACCGCAAGTTC 484
Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 485 ACCCTTCGGATTCTGTACGACCAGGCGAGGGCGGCCAGCCTCTGGTACCCAGCAGCAGG 544
Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 545 CTGAATGGTGTCAACGTGACATATCTCCCTGGGGTTTACATTGCTGGCATCCAGAGGGGC 604
Qy 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 605 ATCATGTCTGAAAGAATGGAATACGACAGGCGGGCCGCGCATCATCCAGGATCTTCGCT 664
Qy 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 665 GATGGAAAGACATGGAGCTACACATACTTAGAGRAGTCCATGTTGTGCTACTACACAGC 724
Qy 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 725 CAGAGGCAGTATATCTTTGAGTTCGACAAGAATGACCCGCCCTCTCTTCTGTGACGATGCC 784
Qy 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 785 AACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT 844
Qy 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 845 CAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGGCACCTCCTT 904
Qy 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 905 CACACCTTCTACCTGGGCACTGGCCGAGGGTGATATACAAGTATGGCAACTGTCAAAG 964
Qy 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 965 CTGGCAGAGACGTTCTATGACACCAAGGTCAAGTTCACCTATGACGAGACGGCAGGC 1024

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 1025 ATGCTGAAGACCACATCAACCTACAGAAATGAGGGCTTCACCTGCACCATCCGCTACCGTCAG 1084
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGlyMetValAsnAla 320
Db 1085 ATTGGGCCCTTGATTGACCGACAGATCTTCGGCTTCACAGAGGATGGTCAACGCC 1144
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 1145 CGTTTGACTACAACATATGACACAGCTTCGGGTGACCAAGCATGCAGGCTGTGATCAAC 1204
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db 1205 GAGACCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGAGCAAGACAGAGCAG 1264
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 1265 TTTGGGAAGTTTGGTGTCAATTACTATGACATTAACAGATCATCAACAGCTGTGTCATG 1324
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 1325 ACCACACCAAGCATTTTGTATGATATGGCAGGATGAAGGATGAGTATGATATGATCTTC 1384
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 1385 CGCTCGCTCATGTACTGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAAG 1444
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 1445 GAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCCTATGATGATGCTGAC 1504
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db 1505 GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGTACGACCTC 1564
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 1565 AATGGGAACCTGCACCTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTAT 1624
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 1625 GACATCCGCGACCGCATCACTCGGCTGGTGACCTGCAATACAAGATGGATGAGATGGC 1684
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 1685 TTCCTGAGGCGAGCGGGCGGTGATATCTTTGAGTACAACACTCAGTGGCTGTCTCATCAAG 1744
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 1745 GCCTACAACCGGGCTGGCAGCTGGAGTGTGAGTACCGTACGATGGCTTGGGGCGGCGC 1804
QY 541 Val 541
Db 1805 GTG 1807
RESULT 8
AX675551 8438 bp DNA linear PAT 27-MAR-2003
LOCUS AX675551
DEFINITION Sequence 1 from Patent WO02055704.
ACCESSION AX675551
VERSION AX675551.1 GI:29333552
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S.,
Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M.,
Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X.,
Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V.,
Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J.,

Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and
Stone,D.J.
Proteins, polynucleotides encoding them and methods of using the
same
JOURNAL Patent: WO 02055704-A 1 18-JUL-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..8438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: Length: 8438
Pred. No.: 4.2e-234 Matches: 541
Score: 2821.50
Percent Similarity: 98.72% Conservative: 0
Best Local Similarity: 98.72% Mismatches: 0
Query Match: 99.17% Indels: 7
DB: 6 Gaps: 2
US-10-029-020-14_COPY_1760_2300 (1-541) x AX675551 (1-8438)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5318 TACTACATCGGGCCGATGGCTCCTTGGCGTGTCTGGCAACGGCATGGAGGTGGCG 5377
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5378 CTGCAGACTGAGCCCCACTTGTGGTGGCACCGTCAACCCCAACCGTGGCAAGAGGAAT 5437
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5438 GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAG 5497
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArg-----ValHisAsnArg 77
Db 5498 GCTCGGGGCCAGGTCACTGTCTTTGGCGCCGCTGCTCCAGGTTCAACACCGA 5557
QY 78 AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHis 97
Db 5558 AATCTCTATCTCTGACTTTGATCGCGTAACACGCACAGAGAAGATCTATGATGACCAC 5617
QY 98 ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerPro 117
Db 5618 CGCAAGTTCAACCTTCGGATTCTGTACGACCATCTCCCTGGGGGTTACATTGTGGCATCC 5677
QY 118 SerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIle 137
Db 5678 AGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTGGGGGTTACATTGTGGCATC 5737
QY 138 GlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArg 157
Db 5738 CAGAGGGGCATCATGCTGAAAAGATGGAATACGACCGCGCGCCGCATCATCCAGG 5797
QY 158 IlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLys-----Ser 173
Db 5798 ATCTTCGCTGATGGGAAGACATGGAGCTACACATATTAGAGAAGGCAGGTGTCCAGTCC 5857
QY 174 MetValLeuLeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArg 193
Db 5858 ATGGTGTGCTACTACACAGCCAGAGGCAGTATATCTTTGAGTTCGACAAGAATGACCCG 5917
QY 194 LeuSerSerValThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerVal 213
Db 5918 CTCTCTTCTGTGACGATGCCCAACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTG 5977
QY 214 GlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPhe 233
Db 5978 GGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACTTC 6037
QY 234 ThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyr 253

Db 6038 ACTGAGGATGGGCACCTCCTTCACACCTTCTACCTGGGCACCTGGCCGCAGGGTGATATAC 6097

QY 254 LysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPhe 273

Db 6098 AAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGTTCAGTTTC 6157

QY 274 ThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThr 293

Db 6158 ACCTATGACGAGACGGCAGGCATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACC 6217

QY 294 CysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThr 313

Db 6218 TGCACCATCCGCTACCGTCAGATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCAC 6277

QY 314 GluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThr 333

Db 6278 GAGGAAGGCATGGTCAACGCCGCTTTTGACTACAACTATGACAAACAGCTTCCGGGTGACC 6337

QY 334 SerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAsp 353

Db 6338 AGCATGCAGGCTGTGATCAACGAGACCCCACTGCCCAATGATCTCTATCGCTATGATG 6397

QY 354 ValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGln 373

Db 6398 GTGTCAGGCAAGACAGACGAGTGTGGGAAGTTTGGTGTCATTTACTATGACATTAAACG 6457

QY 374 IleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLys 393

Db 6458 ATCATCACCAACAGCTGTCATGACCCACACCAAGCATTTTGATGCATATGGCAGGATGA 6517

QY 394 GluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsn 413

Db 6518 GAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATA 6577

QY 414 MetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyr 433

Db 6578 ATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCG 6637

QY 434 SerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeu 453

Db 6638 TCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCATGACAAGCCACT 6697

QY 454 TrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAla 473

Db 6698 TGGCGCTACAGCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTG 6757

QY 474 ArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGln 493

Db 6758 CGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCA 6817

QY 494 TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn 513

Db 6818 TACAAGATGGATGAGGATGGCTTCTCTGAGGCAGCGGGCGGTGATATCTTTGAGTACA 6877

QY 514 SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArg 533

Db 6878 TCAGCTGGCCTGCTCATCAAGGCCCTACAACCGGCTGGCAGCTGGAGTGTGAGGTACCG 6937

QY 534 TyrAspGlyLeuGlyArgArgVal 541

Db 6938 TACGATGGCCTGGGGCGGCGCGTG 6961

RESULT 9

AKI22490

LOCUS AKI22490 5583 bp mRNA linear ROD 15-MAR-2003

DEFINITION Mus musculus mRNA for mKIAA1302 protein.

ACCESSION AKI22490

VERSION AKI22490.1 GI:28972711

KEYWORDS FLI_CDNA.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

CDS

1 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S., Nakajima,D., Nagase,T., Ohara,O. and Koga,H.
Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries
DNA Res. 10, 35-48 (2003)
2 (bases 1 to 5583)
Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
Direct Submission
Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.
Location/Qualifiers
1..5583
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg04788"
/tissue_type="brain"
/dev_stage="adult"
/note="vector:modified pBC SK+"
1..5583
/gene="mKIAA1302"
<1..3597
/gene="mKIAA1302"
/note="CDS is predicted by in silico analysis. Start codon is not identified."
/codon_start=1
/evidence=not experimental
/product="mKIAA1302 protein"
/protein_id="BAC65772.1"
/db_xref="GI:28972712"
/translation="TQNMYELSSPIDQELYLFDTSGKHLYTQSLPTGDYLYNFYTYTGD
GDI THITDNNGNMNVNRDSTGMPLWLVPDQGVYVWMTGNSALRSVTTQGHELAMM
TYHNSGLLATKSNENGWTFEYDSFGRLTNVTFPTQVSSFRSDTSSVHVQVETS
SKDDVTITNLSASGAFYLLQDQVRNSYI GADGSLRLLLANGMEVALQTEPHLLAG
TVNPTVGKRNVTLPIDNGLNVEWRQKEQARQVTFGRRLRVHNRNLLSLDFRVT
RTEKIYDDHRKFTLRILYDQGRPSLWSPSSRLNGVNVYSPGGHIAGIQRGIMSERM
EYDQAGRITSRIFADGKMWSYTYLEKSMVLHLSQRYIFEFDKNDRLSSVTMPNVAR
QTLTIRSVGYRNIYOPPEGNASVIQDFTEDGHLHFTYLGTRRRVIYKYGKLSKLA
ETLYDTKVSFTYDETAGMLKTNLQNEGFTCTIRYQIGPLIDRQIFRTEEGMVNA
RFDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGFVIYDINQIITTA
VMTHKHFDAYGRMKVEQYEIFRSLMYMTVQYDNNGRVVKELKVGYPYANTTRYSYE
YDADGQLQTVSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDLDRITRLGDVQY
KMEDEGFLRQGGDFEYNSAGLLIKAYNRASGVSRYRYDGLGRRVSSKSSHSHLQ
FFYADLTNPTKVTHLYNHSSEITSLYYDLOHUFAMELSSGDEFYIACDNIGTPLAV
FSGTGLMIKQILYAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGR
WTS PDHELWKRLSSNSIVFHLMYFNKNPNISQDIKCFMTDVNSWLLTTFGQLHNV
IPGYKPDTDAMEPSYELVHTQMTQEWDNKSLIGVQCEVQKQKAFVTLERFDQLY
GSTITSCQQAPETKKFASSGSIFGKGVKFAKDKGRVTTDIIISVANEDGRRIAAILNNA
HYLENLHFTIDGVDTHYFVKPGPSEGDLAILGLSGGRRTLENGVNVTVSQINTMLSGR
TRRYTDIQLQYRALCLNTRYGTTVDEKRVLELARQAVRQAWAREQQORLREGEGEL
RAWTDGEKQQVLNTRGVQYDGFVTSVEQYPELSDSANNIHFMRQSEMGRR"

ORIGIN

Alignment Scores:

Pred. No.: 1.33e-233 Length: 5583

Score: 2813.00 Matches: 533

Percent Similarity: 99.26% Conservative: 4

Best Local Similarity: 98.52% Mismatches: 4

Query Match: 98.88% Indels: 0

DB: 10 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AKI22490 (1-5583)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db	1645	TTTGGGAAGTTTGGTGTCTATCTACTACGACATCAACACGATCATTACCACAGCCGTCATG	1704
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
Db	1705	ACCCACACCAAGCACTTTGATGCTTATGGCAGGATGAAGAAAGTACAGTATGAGATTTTC	1764
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420
Db	1765	CGGTCACTCATGTACTGTGATGACTGTTTCACTATGATATGATGATGATGATGATGATGAT	1824
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
Db	1825	GAGCTGAAGGTGGGACCCCTATGTCACACACTACCCGCTACTCTCTATGATGATGATGATGAT	1884
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu	460
Db	1885	GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTATGACCTC	1944
QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
Db	1945	AATGGGAACCTACACTTGTGAGCCCTGGGAACAGCGCAGCGCTACACCACTACGGTAT	2004
QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
Db	2005	GACCTCCGTGACCGCATCACTAGGCTGGTGGTGATGTCTTCGAGTACAATTACGCCGGCTGCTCATCAA	2064
QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys	520
Db	2065	TTCTCTGAGGCAGCGGGTGGGATGTCTTCGAGTACAATTACGCCGGCTGCTCATCAA	2124
QY	521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
Db	2125	GCCTACAACCGGGCTAGTGGTGGAGTGTCAAGTACCGATACCGATGGGCTGGGACGCCGA	2184
QY	541	Val 541	
Db	2185	GTA 2187	
RESULT 10			
AB025413			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
gene			

CDS

```
188. .8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"
/translation="MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKGPQKYSSET
LKAYDQARLAYGSRVMDMPQEAEEFCRTGTFNFTLRELGLGEMTPPHGTLVYRTDIGL
PHCYSGMGASSDADLEADTVLSPHEPVLWGRSTRGRSSCLSSRANSNLTLTPTTEHE
NTETDHPSSLQNHPRLRTPPPPLPHAHTPNQHHAAINSLRNGFTPRSNLPSAPTDH
SLSGEPPAGSAQEPTHAQDNWLLNSNIPLETNLGKQPFGLTLQDNLIEMDILSASRH
DGAYSDDGHFLFKPGTSPLEFCTTSPGYPLTSSTVYSPPRPLRSTFSRPAFNLKPS
KYCNWKCAALSAILISATLIVLLAYFVAMHLFGLNWHLQPMEGQMOMEITETDASSW
PVPTDVSLYPSGGTGLETPDRKGGAAGKPSLSLFPEDSFIDSGEIDVGRRASQKIPP
GTFWRSQVFIHDPVHLKENVSLGKAALVGIYGRKGLPPSHTQFDFVELDGRRLLTQE
ARSLGEPQQRSGRPVPPSSSHETGFIQYLDSGIWHLAFYNDGKESEVVSFLTIAESVD
NCPSNICYNGDCCISGTCHCPLGFLGPDCCRASCPVLCSGNGQYMKGRCLCHSGWKGA
CDVPTNQCIDVACSSHGTCTIMGTICINPGYKGESCEEVDCMDPTCSSRGVCVRGECH
SVGWGNTCETPRATCLDQCSGHGTFPLDTGLCNCDPSTWTHDCSIEICAAADCGHGV
CVGCTCRCEGWMGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHCTIAHYLDRVV
KEGCPGLCNGNRCGLDNLGNWHVCQLGWRGTGCDTSMETGCGDGKNDGDGLVDCMD
PDCCLOPLCHVNPLCLGSPDPLDIIQETQAPVQQNLNPFYDRIKFLVRDSTHSIPG
ENPFDGGHACVIRGQVMTSDGTPLVGVNISFINNPLFGYTI SRQDGSFDLVTNGGISI
ILRFERAPFITQEHTLWLPWDRFFVFMETIVWRHEENEIPSCDLSNFARPNPWWSPSL
TFSASSCAEKGPIVPEIQALQEEIVIAGCKMRLSYLSRTPGYKSVLRISLTHPTIPF
NLMKVHLMVAVEGRLEFKWFAAAPDLSYFVWDKTDVYNQKVFGEFSAFVSVGYEYES
CPDLILWEKRTAVLQYIEIDASKLGWSLDKHALNIQSGILHKNGENQVSPQPPV
IGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCSGDSLYVGDPNYIRRI FPSGNV
TNILEMRNKDPRHSHSPAHHKYLLATDPMGSAVFLSDTNSRRRVFKVKS TTVVKDLVKN
EVVAGTGDOCLDFDTRCGDGKATEATLTPRGITVDKFGLIYFVDGTMIIRVDQNG
IISTLGSNDLT SARPLCDSVMEISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISE
NHQVRIVAGRPMCHQVPGIDHFLLSKVAIHATLESATALAVSHNGVLVIAETDEKKN
RIRQVTTSGEISLVAGAPSGCDCKNDANCDPFGDDGYAKDAKLNTPSSLAVCAGEL
YVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLEFDTSGKHYLTSQSLPTGDYLY
NFTYTGDDIITHITDNGNMVNVRDSTGMPLWLVPDGVVWVTMTNSALRSVTTQ
GHELAMTYHGNSGLLATKSNENGWTTTFEYDSFGRLTNVTFTGTQVSSFRSDTSSV
HVQVETSSKDDVTITNLSASGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQT
EPHLLAGTVNPTVGKRNVTLPIDNGLNLVWRQKEARGQTVFGRRLRVHNRNLS
LDFRVTREKIYDDHRKFTLRILYDQAGRPSLWSPSRLNGVNVTVSPGGHAGIOR
GIMSERMEYDQAGRITSRIFADGKMWSTYILEKSMVHLHSQRQYI FFDKNDRLSSV
TMNVARQTLTETIRSVGYRNIIYQPPGNASVIQDFTEDGHLHHTFYLGTGRRVIYKY
GKLSKLAETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCTIRYQIGPLIDRQIFRFT
EEGMVNARFDYNDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYDI
NQIITTAVMTHTKHFDAYGRMKEVQYEIFRSLMYMWTQYDNMGRVVKELKVGYPAN
TRYSEYDADGOLQTVSINDKPLWYSYDLNGLNHLSPGNSARLTPLRYDLDRIT
PLGDVQYKMDDEGFLRQGGDVFEYNSAGLLIKAYNRASGWSVRYRYDGLGRRVSXS
SHSHLQFFFYADLTNPTKVTHLYNHSSEITSLYYDLQGHLFAMELSSGDEFYIACDN
IGTPLAVFSGTGLMIKQLIYATYGEIYMDTNPNFQIIGYHGGLYDPLTKLVEMGRD
YDVLAGRWTSPDHELWKRLSSNSIVPFHLYMFKNNPIINSQDIKCFMTDVSNNLLTF
GFQLHNVI PGYPKPDTDAMEPSYELVHTQMTKEWDNDSKSI LGVQCEVQKQKAFVTL
ERFDQLYGSTITSQQAPETKKFASGSI IFGKGVKFALKDGRVTTDIIISVANEGRRI
AAILNNAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGRRRTLENGVNVTVSQI
NTWLSGRTRRYTDIQLQYRALCLNTRYGTTVDEEKRVVLELARQRAVRQAWAREQORL
REGEGLRAWTDGKEKQVVLNTRVQYDGFVTSVEQYPELSDSANNIHFMRQSEMGR
R"
```

ORIGIN

Alignment Scores:
Pred. No.: 2.34e-233 Length: 8585
Score: 2813.00 Matches: 533
Percent Similarity: 99.26% Conservative: 4
Best Local Similarity: 98.52% Mismatches: 4
Query Match: 98.88% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AB025413 (1-8585)

QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db	5471	TACTACATCGGGCTGATGGCTCCCTGAGGCTGCTACTAGCCAATGGCATGGAAGTGGCT 5530
QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluThrArgGlnArgLysGluGln 60
Db 6148 GTCACACTGCCCATTTGATAATGGCCCTCAACCTGGTGGAGTGGCGCAGCGCAAGGAGCAG 6207
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 6208 GCTCGTGGCCAGGTCACCGTCTTTTGGACCCCGTYTGGCGGTTCAACCCGAAACCTCTTG 6267
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 6268 TCTTTGGACTTTTGACCGTGTCAACGACAGAGAGATCTACGATGACCATCGCAAGTTC 6327
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrTrpSerProSerSerArg 120
Db 6328 ACCCTTCGGATCCTATATGACCGCAGGAGGCCCGCCAGCTTCTGGTCACTAGCAGCAGG 6387
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 6388 CTGAATGGTGTTAATGTGACCTACTCCCTGGAGGTACATTGCTGGAATCCAAAGGGGC 6447
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 6448 ATCATGTCTGAGAGAAATGAATAATGATCAGGCAGGCGCATCACATCCCGGATCTTTGCA 6507
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
Db 6508 GACGGGAAAATGTGGAGCTACACGTACTTAGAGAAAGTCCATGGTGTCTTCACTCTCCACAGC 6567
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 6568 CAGAGGCAGTACATCTTCAGTTTGACAAAGATGACCGCTCTCTTCTGTGACCATGCC 6627
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 6628 AACGTAGCCCGGCAGACGCTGGAGACCATCCGCTCAGTGGGTACTACAGGAACATCTAC 6687
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 6688 CAGCCCCCGGAAGGCAACGCCTCAGTCATTACGACTTCACTGAGGATGGACACCTGTTA 6747
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6748 CATACTTCTACTTTGGGCACCGGCCCGGGTGATTTACAAGTATGGCAAGTGTCAAAG 6807
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6808 CTGGCCGAGACTCTGTATGACACCACCTAAGSPAAGCTTCACCTACGACGAGACCGCAGG 6867
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6868 ATGCTGAAGACTGTCAACCTACAGAATGAGGCTTCACCTGCACACTCCGCTACCGTCAG 6927
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6928 ATTGGGCCCTTGATTGATCGGCAGATCTTCGTTTCACGGAAGAAGCATGGTCAATGCC 6987
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6988 CGTTTTGATTACAACACTATGACAAACAGTTTCCGTGTGACTAGCATGCAAGGCTGTGATCAAT 7047
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 7048 GAGACCCCACTGCCCATTTGACCTCTACCGCTATGATGATGTGTGAGGAAGACAGAGCAG 7107
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 7108 TTTGGGAAGTTTGGTGTCAATTTACTACACATCAACCAGATCATTACCACAGCCGTCATG 7167
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 7168 ACCCACTCCAAGCACTTTGATGCTTATGGCAGGATGAAGGAAGTACAGTATGAGACTTTC 7227

QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLysLys 420
Db 7228 CGGTCACCTCATGTACTGGATGACTGTTTCAGTATGATAACATGGGACGGGTAGTGAAGAAG 7287
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 7288 GAGCTGAAGGTGGGACCCCTATGCCAACACTACCCGCTACTCCTATGAGTATGATGCTGAT 7347
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db 7348 GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTATGACCTC 7407
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 7408 AATGGGAACCTACACTTGTCTGAGCCCTGGGAACAGCGCAGCTCACACCACTACGGTAT 7467
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 7468 GACCTCCGTGACCGCATCACTAGGCTGGGTGATGTACAGTACAAGATGGATGAGGATGGC 7527
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 7528 TCCCTGAGGCAGCGGGTGGGATGTCTTCGAGTACAAATTCAGCCGGGCTGCTCATCAA 7587
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 7588 GCCTACAATCGGGCTAGTGGGTGGAGTGTTCAGGTACCGATACGATGGGCTGGGACGCCGA 7647
QY 541 Val 541
Db 7648 GTA 7650
RESULT 12
AB026980
LOCUS
DEFINITION
AB026980
ACCESSION
AB026980.1 GI:5307784
VERSION
KEYWORDS
ten-m4.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (sites)
AUTHORS
Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.
TITLE
Compartmentalized expression of zebrafish ten-m3 and ten-m4,
homologues of the Drosophila ten(m)/odd Oz gene, in the central
nervous system
JOURNAL
Mech. Dev. 87 (1-2), 223-227 (1999)
MEDLINE
99425191
PUBMED
10495292
REFERENCE
2 (bases 1 to 9264)
AUTHORS
Mieda,M.
JOURNAL
Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute,
RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hirosawa,
Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp,
Tel:81-48-467-9713, Fax:81-48-467-9714)
COMMENT
Sequence updated (29-Jun-1999).
FEATURES
Location/Qualifiers
1. .9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1. .9264
/gene="ten-m4"
323. .8797
/gene="ten-m4"
/note="similar to Drosophila melanogaster tenm/odx and Mus
musculus Doc4"
/codon_start=1

Db 6965 CGTTCGCTCATGTATTGGATGATGGTACAGTTTGACAACATGGGGCGAGTGGTGGCCAAG 7024

QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaasp 440

Db 7025 GAGCTTAAAGTCGACCCCTATGCAAAATACCACCCGCTATGCCTATGAATATGATGCAGAT 7084

QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460

Db 7085 GGTCACTCCAGGTAGTGTCATCAATGATAAACCTCTATGGCGCTACAGTTATGACCTT 7144

QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480

Db 7145 AATGGCAATTGCACTCTTAGCCCTGGTAATAGCGCAGACTTACACCACTACGCTAC 7204

QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500

Db 7205 GACATTCGGGACCGCATTACTCGTCTAGGTGATGTGCAGTACCGTTTAGATGAGGATGGA 7264

QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520

Db 7265 TTCTCTCAGGCAAGAGGCAATGATTTCTTTGAGTATAACTCTGCTGGTTCGTTGTTAA 7324

QY 521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540

Db 7325 ACCTACAACAAGTTAACGGATGGACTATCAAGTACCGCTATGATGGCTTGGCAGAAGA 7384

QY 541 Val 541

Db 7385 GTG 7387

RESULT 13

AX876525

LOCUS AX876525 3614 bp DNA linear PAT 17-DEC-2003

DEFINITION Sequence 11430 from Patent EP1074617.

ACCESSION AX876525

VERSION AX876525.1 GI:40031261

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

TITLE Primers for synthesising full-length cDNA and their use

JOURNAL Patent: EP 1074617-A 11430 07-FEB-2001; Research Association for Biotechnology (JP)

FEATURES

Location/Qualifiers

1..3614

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

149..3286

/note="unnamed protein product"

/codon_start=1

/protein_id="CAE89315.1"

/db_xref="GI:40031262"

/translation="MDKAITVDIESSSREEDVSITSNLSSIDSFYTMVQDLRNSYQI

GYDGLSLRIIYASGLDSHYOTEPHVLAGTANPTVAKRNTLPGENGQNLVEWFRKEQA

QGKVNVFGRKLRVNGRNLLSVDFDRTTKTEKIYDHRKFLRLRIAYDTSGHPTLWLPSS

KLMAVNVITYSSTGQIASIORGTSEKVDYDQGGRIVSRVFADGKTWSYTYLEKSMVLL

LHSRQYIPEYDMWRDLSATMPSVARHTMQTIRSIGYRNIYNPPESNASIITDYNE

EGLLQTAFLGTSRRVLFKYRRQTRLSEILYDSTRVSFTYDETAGVLKTVNLQSDGFI

CTIRYRQIGPLIDRQIFPFSDEGMVNARFDYSYDNSFRVTSMQGVINETPLPIDLYQF

DDISGKVEQFGKFGVIYDINQIISTAVMTYTKHDAHGRIKEIQEIEFRSLMWITI

QYDNMGRVTKREIKIGFPANTKYAYEYDVQGLQTVLNEKIMWRNYNDLNGNLHL

NPSNARLTPLRYDLDRITRLGDVQYRLDEDEFLRQRGTEIFEYSKGLLFRVYSKG

SGWTVIYRDGLGRRVSSKTSLGHLQFFYADLTYPTRITHVYNHSSSEITSLYYDLQ

GHLFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLKQIQYTAYGEIYFDSNIDFQLVI

GFGHGLYDPLTKLIHFGERDYDILAGRWTTPDIEWKRIKDPAPFNLYMFRNNPAS

KIHDVKDYITDVNSWLVTFGFHLHNAIPGFPVPKFDLTPESYELVKSQQWDDIPPIFG

VQQVARQAKAFLSLGKMAEVQVSRRRAGGAQSWLWFATVKSLLIGKVMLAVSQGRVQ

TNVLNIANEDCIKVAAVLNNAFYLENLHFTIEGKDTHYFIKTTTPESDLGLRLTSGR

ORIGIN

Alignment Scores:

Pred. No.: 1.42e-183 Length: 3614

Score: 2233.00 Matches: 404

Percent Similarity: 88.91% Conservative: 77

Best Local Similarity: 74.68% Mismatches: 60

Query Match: 78.49% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AX876525 (1-3614)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20

Db 272 TACCAGATTGGTTATGACGGCTCCCTCAGAAATTATCTAGCCAGTGGCCTGGACTCACAC 331

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 332 TACCAAAACAGAGCCGACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAAAC 391

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60

Db 392 ATGACTTTGCCCTGGCGAGAAACGGTCAAAACTTGGTGGAAATGGAGATTCCGAAAAAGAGCAA 451

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80

Db 452 GCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTT 511

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100

Db 512 TCAGTTGACTTTGATCGAACAAACAAGACAGAAAAAGATCTATGACGACCAACCGTAAATT 571

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerSerArg 120

Db 572 CTACTGAGGATCGCTACGACACGTCCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAG 631

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140

Db 632 CTGATGGCCGTCATGTACCTATTTCATCCACAGGTCAAATTGCCAGCATCCAGCGAGGC 691

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 692 ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTCTCGGGTCTTTGCT 751

QY 161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180

Db 752 GATGTTAAAAACATGGAGTTACACATATTAGAAAAGTCCATGGTTCCTTCTGCTTCATAGC 811

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

Db 812 CAGCGCAGTACATCTTCGAATACGATATGTGGGACCGCTGTCTGCCATCACCATGCC 871

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 872 AGTGTGGCTGCCACACCATGAGACCATCCGATCCATTGGTACTACCGCAACATATAC 931

QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 932 AACCCCGGAAAGCAACGCCCTCCATCATCAGGACTACAAACGAGGAAGGGCTGCTTCTA 991

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260

Db 992 CAAACAGCTTCTTGGGTACAAGTCGAGGGTCTTATTCAAATACAGAAGGAGACTAGG 1051

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 1052 CTCTCAGAAATTTTATATGATAGCACAGAGTCAGTTTACCTATGATGAACACAGCAGGA 1111

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db	1112	GTCCTAAAGACAGTAAACCTCCAGAGTGATGGTTTTATTTCACCATTAGATACAGGCAA	1171
QY	301	IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla	320
Db	1172	ATTGGTCCCTGATTGACAGGCAGATTTTCGCCTTTAGTGAAGATGGGATGGTAAATGCA	1231
QY	321	ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn	340
Db	1232	AGATTGACTATAGCTATGACAAACAGCTTTCGAGTGACCAGCATGCAGGGTGTGATCAAT	1291
QY	341	GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln	360
Db	1292	GAACGCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG	1351
QY	361	PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrAlaValMet	380
Db	1352	TTTGGAAGATTGGAGTTATATATTATGATATTAACAGATCATTTCTACAGCTGTAATG	1411
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
Db	1412	ACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC	1471
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLysLys	420
Db	1472	AGGTGCTCATGTACTGGATTACAATTCAAGTATGATAACATGGTGGGTAAACCAAGAGA	1531
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
Db	1532	GAGATTAAATAATAGGCCCTTTTGCCAACACCAACCAATAATGCTTATGAATATGATTTGAT	1591
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu	460
Db	1592	GGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAAATGTGGCGGTACAACTACGATCTG	1651
QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
Db	1652	AATGGAAACCTCCATTACTGAACCAAGTAACAGTGCAGTCTGACACCCCTTCGCTAT	1711
QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
Db	1712	GACCTGGAGACAGAAATCACTCGACTGGGTGATGTTCATATCGGTTGGATGAAGATGGT	1771
QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
Db	1772	TTCTTACGTCAAAGGGCAGCGAAATCTTTGAATATAGCTCCAAGGGCTTTAACTCGA	1831
QY	521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
Db	1832	GTTTACAGTAAAGGAGTGGCTGGACAGTGTCTACCGTTATGACGGCCTGGGAAGCGGT	1891
QY	541	Val	541
Db	1892	GTT	1894
RESULT 14			
BD156175			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

3614 bp

DNA

linear

PAT 17-JAN-2003

BD156175

Primer for synthesizing full-length cDNA and use thereof.

BD156175

BD156175.1

GI:27861933

JP 2002191363-A/11018.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3614)

Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A.11018 09-JUL-2002;

HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/11018

PD	09-JUL-2002	
PF	28-JUL-2000	JP 2000280990
PI	TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO,	
PI	JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, KEIICHI NAGAI, TETSUJI OTSUKI	
PC		
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC		
PC	Primer for synthesizing full-length cDNA and use thereof	Key
	Location/Qualifiers	
FT	CDS	(149).. (3283).
FEATURES	Location/Qualifiers	
	source	
	1..3614	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
Alignment Scores:		
Pred. No.:	1.42e-183	Length: 3614
Score:	2233.00	Matches: 404
Percent Similarity:	88.91%	Conservative: 77
Best Local Similarity:	74.68%	Mismatches: 60
Query Match:	78.49%	Indels: 0
DB:	6	Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x BD156175 (1-3614)		
QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db	272	TACCAGATTGGTTATGACGGCTCCCTCAGATTATCTACGCCAGTGGCCTGGACTCACAC 331
QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db	332	TACCAACACAGAGCCGACGTTCTGGCTGGCACCCTAATCCGACGGTTCGCAAAAGAAAC 391
QY	41	ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
Db	392	ATGACTTTGCCGAGAACGGTCAAAACCTGGTGGATGGAGATTCCGAAAGAGCAA 451
QY	61	AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db	452	GCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGTTAATGGCAGAAACCTCCTT 511
QY	81	SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db	512	TCAGTTGACTTTGATCGAACAAACAAAGACAGAAAGATCTATGACGACCCGTAATTT 571
QY	101	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db	572	CTACTGAGGATCGCCTACGACACGCTCTGGGCACCCGACTCTCTGGTGCACAGCAAG 631
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db	632	CTGATGGCCGTCATGTACCTATTTCATCCACAGGTCAAAATGCCAGCATCCAGCGAGGC 691
QY	141	IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db	692	ACCAGTGGAGAAAGTAGATTATGACGGACAGGGGGAGGATCGTGTCTCGGTCCTTGCT 751
QY	161	AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db	752	GATGGTAAACATGGAGTTACACATATTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 811
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db	812	CAGCGCAGTACATCTTCGATACGATATGTGGACCGCCTGTCTGCCATCACCATGCC 871
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db	872	AGTGTGGCTCGCCACACACCATGCAGACCATCCGATCCATTGGCTACTACCGCAACATATAC 931

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 932 AACCCCGGAAAGCAACGCCTCCATCATCACGGACTACAACGAGGAGGGCTGCTTCTA 991
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 992 CAAACAGCTTCTTGGGTACAAAGTCGGAGGGTCTTATTCAATACAGAAGGCAGACTAGG 1051
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 1052 CTCTCAGAAATTTTATATGATAGCACAAAGAGTCAGTTTACCTATGATGAACACAGCAGGA 1111
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 1112 GTCTAAAGACAGTAAACCTCCAGAGTGATGGTTTATTTCACCATTTAGATACAGGCAA 1171
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 1172 ATTGGTCCCTGATTGACAGGCAGATTTTCGCTTTAGTGAAGATGGGATGGTAAATGCA 1231
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 1232 AGATTGTACTATAGCTATGACACACAGCTTTCGAGTGACCATGTCAGGGTGTGATCAAT 1291
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db 1292 GAAACGCCACTGCCATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG 1351
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 1352 TTTGGAAAGTTTGGAGTTATATATTATGATATTAAACCAGATCATTTCTACAGCTGTAATG 1411
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 1412 ACCTATACGAAGACACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC 1471
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 1472 AGGTCGCTCATGTACTGATGATTACAAATTCAGTATGATAACATGGGTCCGGTAACCAAGAGA 1531
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 1532 GAGATTAAATAGGGCCCTTTGCCAACACCAACCAATATGCTTATGAATATGATGTTGAT 1591
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db 1592 GGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACTACGATCTG 1651
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 1652 AATGAAACCTCCATTTACTGAACCAAGTAACAGTGCAGTGCAGTGCACACCCCTTCGCTAT 1711
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 1712 GACCTGCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT 1771
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 1772 TTCCTACGTCAAAGGGCCACGGAATCTTTGAATATAGCTCCAAGGGGCTTCTAACTCGA 1831
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 1832 GTTTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCGTTATGACGGCCTGGGAAGCGCT 1891
QY 541 Val 541
Db 1892 GTT 1894

RESULT 15
AK001336
LOCUS AK001336 3614 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ10474 fis, clone NT2RP2000067.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source

CDS

ORIGIN

Alignment Scores:

Pred. No.: 1.42e-183 Length: 3614
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60

AK001336
AK001336.1 GI:7022530
oligo capping; fis (full insert sequence).
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,
Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,
Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3614)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection:
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
Location/Qualifiers
1. .3614
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP2000067"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 2-weeks retinoic acid (RA)
induction."
149. .3286
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91633.1"
/db_xref="GI:7022531"
/translation="MDKAITVDIESSREEDVSIITNSLSSIDSFYTMVODQLRNSYQI
GYDGLRIIYASGLDSHYQTEPHVLAGTANPTVAKRNTLPGENGQNLVEWFRKEQA
QKKNVFGRLRVNGRNLLSVDFDRTTKTEKIYDHRKFLRLRIAYDTSGHPTLWLPSS
KLMVNVTVSYSTGQIASIQRGTTSEKVDYDQGRIVSRVADGKTWSYTYLEKSMVLL
LHSQRQYIFEYDMWDRLSAITMPSVARHTMQTIRSTIGYVRYNPPESNASIITDYN
EGLLLQTAFLCTSRRLFKYRRQTRLSEILYDSTRVSFTYDETAGVLKTNLQSDGFI
CTIRYQIGPLIDRQIFRSEDGMVNARFDYSYDNSFRVTSMQGVINETPLPLDYQF
DDISGKVEQFGKGVYYDINQIIISTAVMTYTKHFDAGRIKEIQYEIIFRSLMYWITI
QYDNMGRVTKREIKIGPFANTTKYAYEYVDGQLQTVYLNEKIMWRYNYDLNGLHLL
NPSNSARLTPLRYDLDRITRLGDVQVRLDEDFLRQGRTEIFEYSSKGLLTRVYSKG
SGWTVIYRYDGLGRVSSKTSLGQHLQFFYADLTYPTRITHVNHSSSEITSLYDLQ
GHLFAMEISSGDEFYIASDNTGTPLAVSSNGMLKQIQYTAYGEIYFDSNIDFQLVI
GFHGLYDPLTKLIHFGERDYDILAGRWTTPDIEIKRIGKDPAPFNLYMFRNNNPAS
KIHVDKDYITDVNSWLVTFGFHLNAIPGFPVPKFDLTPESELVKVQQWDDIPPIFG
VQQQVARQAKAFLSLGKMAEVQVRRRAGGAQSWLWATVKSILGKVMIAVSGQRVQ
TNVLNIANEDCIKVAALNNAFYLENLHFTIEGKDTHYFKITTTPESLGTLRLTSGR
KALENGINTVQSSTTVVNGTRRRFADVEMQFALALHVRVYGMTLDEEKARILEQARQ
RALARAWAREQQRVDRDGEEGARLWTEGEKRLLSAGKVQGYDGYVLSVEQYPELADS
ANNIQFLROSEIGRR"

Page 20

Query Match:	78.49%	Indels:	0
DB:	9	Gaps:	0
US-10-029-020-14_COPY_1760_2300 (1-541) x AK001336 (1-3614)			
QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla	20
DB	272	TACCAGATTGGTTATGACGGCTCCCTCAGAATTATCTACGCCAGTGGCCTGGACTCACAC	331
QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn	40
DB	332	TACCAAAACAGAGCCGACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAGAAAC	391
QY	41	ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln	60
DB	392	ATGACTTTCCTGGCGAGAACCGTCAAAACTTGGTGAATGGAGATCCGAAAAGAGCAA	451
QY	61	AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu	80
DB	452	GCCCAAGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCT	511
QY	81	SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe	100
DB	512	TCAGTTTGACTTGTATCGAACAAACAGACAGAAAAAGATCTATGACGACCCGTAATTT	571
QY	101	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrProSerSerArg	120
DB	572	CTACTGAGGATCGCCTACGACACGTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAG	631
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly	140
DB	632	CTGATGGCCGTCAAATGTCACCTATTTCATCCACAGGTCAAATTGCCAGCATCCAGCGAGGC	691
QY	141	IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla	160
DB	692	ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGAGGATCGTGTCTCGGTCTTTGCT	751
QY	161	AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer	180
DB	752	GATGGTAAACATGGAGTTACACATATTTAGAAAGTCCATGGTTCTTCIGCTTCATAGC	811
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro	200
DB	812	CAGCGGCAGTACATCTCGAATACGATATGTGGGACCGCTGTCTGCCATCACCATGCC	871
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr	220
DB	872	AGTGTGGCTCGCCACCACTGCAGACCATCCGATCCATGGCTACTACCCGCAACATATAC	931
QY	221	GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu	240
DB	932	AACCCCCCGAAAGCAACCGCTCCATCATCACGGACTACAAACGAGGAGGGCTGCTCTA	991
QY	241	HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys	260
DB	992	CAAACAGCTTCTTGGGTACAAAGTCGGAGGGTCTTTAATCAAAATACAGAAGGCAGACTAGG	1051
QY	261	LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly	280
DB	1052	CTCTCAGAAAATTTTATATGATAGCAACAAGATCAGTTTTTACCTATGATGAACACAGCAGGA	1111
QY	281	MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln	300
DB	1112	GTCCATAAGACAGTAAACCTCCAGAGTGATGGTTTTTATTGACCACTTAGATACAGGCCAA	1171
QY	301	IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla	320
DB	1172	ATTGGTCCCTGATTGACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGGTAAATGCA	1231
QY	321	ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn	340
DB	1232	AGATTTGACTATAGCTATGACAACAGCTTTCGAGTGACCCAGCATGCAGGGTGTGATCAAT	1291

QY	341	GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln	360
DB	1292	GAAACGCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG	1351
QY	361	PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet	380
DB	1352	TTTGAAAAGTTTGGAGTTATATATATGATATTAACACAGATCATTTCTACAGCTGTATG	1411
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
DB	1412	ACCTATACGAAGCACCTTTGATGCTCATGGCCGTATCAAGGAGATTCATATGAGATATTC	1471
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420
DB	1472	AGGTCGCTCATGTACTGGATTACAATTCAGTATGATAACATGGGTCCGGTAAACCAAGAGA	1531
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
DB	1532	GAGATTAAAATAGGGCCCTTTGCCAACACCCAAATATGCTTATGAATATGATGTTGAT	1591
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu	460
DB	1592	GGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACTACGATCTG	1651
QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
DB	1652	AATGAAACCTCCATTTACTGAACCCCAAGTAACAGTGGCGTCTGACACCCCTTCGCTAT	1711
QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
DB	1712	GACCTGCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT	1771
QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuIleLys	520
DB	1772	TTCTTACGTCAAAGGGGCACGGAATCTTTGAATATAGTCTCCAAGGGGCTTCTAACTCGA	1831
QY	521	AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
DB	1832	GTTTACAGTAAAGCCAGTGGCTGGACAGTGATCTACCGTTATGACGGCCTGGGAAGGCGT	1891
QY	541	Val	541
DB	1892	GTT	1894

Search completed: August 14, 2004, 12:02:17
Job time : 9668.56 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 3555.05 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGYHGGGLYDPLTKLVHMGR.....TDIISVANEDGRRVAALNH 201

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_19065@runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1077	100.0	3320	9	AK127101	AK127101 Homo sapi
2	1077	100.0	3486	9	HSM806114	BX537983 Homo sapi
3	1077	100.0	8354	6	AX556500	AX556500 Sequence
4	1077	100.0	8624	9	AB037723	AB037723 Homo sapi
5	1077	100.0	8645	6	AX600210	AX600210 Sequence
6	1077	100.0	8993	9	HSM806812	BX640737 Homo sapi
7	1070	99.4	2685	9	AK056531	AK056531 Homo sapi
8	1063	98.7	8438	6	AX675551	AX675551 Sequence
9	1031	95.7	5583	10	AK122490	AK122490 Mus muscu
10	1031	95.7	8585	10	AB025413	AB025413 Mus muscu
11	1031	95.7	9722	10	AF059485	AF059485 Mus muscu
12	807	74.9	9264	5	AB026980	AB026980 Danio rer
13	667.5	62.0	8118	5	GGA238613	AJ238613 Gallus ga
14	636.5	59.1	8373	10	AB025410	AB025410 Mus muscu
15	630.5	58.5	8297	9	AF100772	AF100772 Homo sapi
16	627	58.2	834	5	GGA293019	AJ293019 Gallus ga
17	566	52.6	2157	6	AX876360	AX876360 Sequence
18	566	52.6	2157	6	BD156088	BD156088 Primer fo
19	566	52.6	2157	9	AK027473	AK027473 Homo sapi
20	566	52.6	3270	6	AX877449	AX877449 Sequence
21	566	52.6	3270	6	BD156663	BD156663 Primer fo
22	566	52.6	3270	9	AK001748	AK001748 Homo sapi
23	566	52.6	3614	6	AX876525	AX876525 Sequence
24	566	52.6	3614	6	BD156175	BD156175 Primer fo
25	566	52.6	3614	9	AK001336	AK001336 Homo sapi
26	566	52.6	5309	9	AB040888	AB040888 Homo sapi
27	566	52.6	8473	6	AX662357	AX662357 Sequence
28	566	52.6	8487	6	AX662359	AX662359 Sequence
29	566	52.6	8645	6	AX662355	AX662355 Sequence
30	566	52.6	8675	6	AX662353	AX662353 Sequence
31	563	52.3	5804	10	AK122513	AK122513 Mus muscu
32	563	52.3	7816	10	AF195418	AF195418 Mus muscu
33	563	52.3	8964	10	AB025412	AB025412 Mus muscu
34	536.5	49.8	3184	10	AF195419	AF195419 Mus muscu
35	536.5	49.8	8797	6	AX250063	AX250063 Sequence
36	536.5	49.8	8797	6	AX250066	AX250066 Sequence
37	536.5	49.8	8797	10	AB025411	AB025411 Mus muscu
38	531.5	49.4	7713	9	HSM808325	BX648178 Homo sapi
39	531.5	49.4	8689	6	AX250067	AX250067 Sequence
40	531.5	49.4	8689	10	AF086607	AF086607 Rattus no
41	529.5	49.2	2926	9	HSM802230	AL137500 Homo sapi
42	529.5	49.2	6560	6	AX250065	AX250065 Sequence
43	529.5	49.2	7781	9	AB032953	AB032953 Homo sapi
44	529.5	49.2	9729	6	AX250013	AX250013 Sequence
45	529.5	49.2	9826	6	AX250008	AX250008 Sequence

ALIGNMENTS

AK127101	LOCUS		3320 bp		mRNA		linear		PRI 09-SEP-2003	
	DEFINITION		Homo sapiens cDNA FLJ45158 fis, clone BRAWH3043034, highly similar to Mus musculus neuregulin 1 (Nr91).							
AK127101	ACCESSION		AK127101.1		GI:34533862					
	VERSION		oligo capping; fis (full insert sequence).							
AK127101	KEYWORDS		Homo sapiens (human)							
	SOURCE		Homo sapiens							
AK127101	ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
	REFERENCE		1		Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.					
AK127101	AUTHORS		NEDO human cDNA sequencing project							
	TITLE		Unpublished							
AK127101	JOURNAL		2		(bases 1 to 3320)					
	REFERENCE		Isogai,T. and Yamamoto,J.							
AK127101	AUTHORS		Direct Submission							
	TITLE		Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan							
AK127101	JOURNAL		(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)							
	COMMENT		NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.							
AK127101	FEATURES		Location/Qualifiers							
	source		1. 3320		/organism="Homo sapiens"					
AK127101			/mol_type="mRNA"							
			/db_xref="taxon:9606"							
AK127101			/clone="BRAWH3043034"							
			/tissue type="brain"							
AK127101			/clone_lib="BRAWH3"							
			/note="cloning vector: pME18SFL3"							
AK127101	ORIGIN									
	Alignment Scores:		1.5e-111		Length:		3320			
AK127101	Pred. No.:		1077.00		Matches:		201			
	Score:		100.00%		Conservative:		0			
AK127101	Percent Similarity:		100.00%		Mismatches:		0			
	Best Local Similarity:		100.00%		Indels:		0			
AK127101	Query Match:		100.00%		Gaps:		0			
	DB:		9							
AK127101	US-10-029-020-14_COPY_2400_2600		(1-201) x AK127101		(1-3320)					
	QY		1		ileGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg		20			
AK127101	Db		1956		ATAGGCTACCATGTGGCTCTATGATCCACTCACAAGCTTGTCCACATGGCGCGCA		2015			
	QY		21		AspTyrAspValLeuAlaGlyArgTTrpThrSerProAspHisGluLeuTrpLysHisLeu		40			
AK127101	Db		2016		GATTATGATGTGCTGGCGGACGCTGGACTAGCCAGACCACGAGCTGTGGAAGCACCTT		2075			
	QY		41		SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer		60			
AK127101	Db		2076		AGTAGCAGCAACGTCATGCCTTTTAATCTCTATATGTTCAAAAACAACAACCCCATCAGC		2135			
	QY		61		AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly		80			
AK127101	Db		2136		AACTCCCAGGACATCAAGTGTCTTCATGACAGATGTTAACAGCTGGCTGCCTCACCCTTGA		2195			
	QY		81		PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro		100			

protein id="CAD97943.1"
/db_xref="GI:31874054"
/translation="SFGRLTNVTFTPTQGVSSFRSDTSSVHVQVETSSKDDVTITTNL
SASGAFYLLQDOVRNSYYIGADGSLRLLLANGMEVALQTEPHLAGTVNPTVGRNV
TLPIDNGLNLEVRQRKEAQGVTVFGRRLRVHNRNLLSLGFDRVTRTEKIYDDHRK
FTLRILYDQAGRPSLWSPSSRLNGVNVTPSPGGYIAGIQIGIMSERMEYDQAGRITSR
IFADGKTWSYTYLEKSMVLLLSORQYIFEDKNDRLSSVTMPNVARQTLETIRSVGY
YRNIYQPEGNASVIQDFTEDGHLLHTPYLGTGRRVIYKYGKLSKLAETFYDTTKVSF
TYDETAGMLKTINLQNEGFTCTIRYQIGPLIDRQIFRFTTEGMVNARFDYNDNSFR
VTSQAVINETPLPIDLRYDDVSGTEQFGKFGVIYDINQIITTAVMTHTKHFDAY
GRMKEVQYEIIFRSLMYMTVQYDNMGRVVKELKVGPIYANTTRYSEYDADGQLQTVS
INDKPLWRYSYDLNGLNHLSPGNSARLITPLRYDIRDRITRLGDVQYKMDDEDGFLRQR
GGDIFEYNSAGLLIKAYNRAGSWSRYRDGLGRRVSSKSSHHLQFFFYADLTNPTK
VTHLVNHSSEITSLYYDLQHLFAMELSSGDEFYIACDNIGTFLAVFSGTGLMIKQI
LYTAGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKH
LSSSNVMPENLYMEKNNPISNSQDIKCFMTDVNSWLLTFGQLHNVIPGYPKPMDMA
MEPSYELHTQMKTOEWDNSKSIILGVQCEVQKQKAFVTLERFDQLYGSTITSQQAP
KTKKFASSGSVFGKGVKALKDGRVTDIISVANEDGRRVAILNHAHYLENLHFTID
GVDTHYFVKPSPSEGD LAILGLSGGRTLENGVNVTVSQINTVLNGRTRRYTDIQLQY
GALCUNTRYGTTLDEEKARVLELARQRAVRQAWAREQQRLREGEGLRAWTEGEKQQV
LSTGRVQGYDGFVISVEQYPELSDSANNIHFMRQSEMGR"
3468
polyA_site
/gene="DKFZp686D0412"

ORIGIN

Alignment Scores:
Pred. No.: 1.59e-111 Length: 3486
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x HSM806114 (1-3486)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 2105 ATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCACATGGCGCGGA 2164
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 2165 GATTATGATGTGCTGGCCGACGCTGGACTAGCCACAGACCATGCTGTGGAACACCTT 2224
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 2225 AGTAGCAGCAACGTGTCATGCCTTTTAATCTCTATATGTTCAAAAACAACAACCCCATCAGC 2284
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 2285 AACTCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTTGA 2344
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 2345 TTCCAGCTACACAACAGTGATCCCTGGTTATCCCAAACCCAGACATGGATGCCATGGAAACC 2404
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 2405 TCCTACGAGCTTATCCACACACAGATGAAAACGCAGAGTGGGACAAACAGCAAGTCTATC 2464
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 2465 CTCGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCTTTGTACCTTAGAACCGTTT 2524
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 2525 GACCAGCTCTATGGCTCCACATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAGTTT 2584
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 2585 GCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTG 2644
QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200

Db 2645 ACCACAGACATCATCAGTGTGTGCCAATGAGGATGGGCGAAGGGTTGCTGCCATCTTGAAC 2704
QY 201 His 201
Db 2705 CAT 2707
RESULT 3
AX556500
LOCUS: AX556500 8354 bp DNA linear PAT 27-NOV-2002
DEFINITION: Sequence 13 from Patent WO02057453.
ACCESSION: AX556500
VERSION: AX556500.1 GI:25899736
KEYWORDS:
SOURCE: Homo sapiens (human)
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS: Gangolli, E.A., Patturajan, M., Vernet, C.A., Malyankar, U.M.,
Kekuda, R., Stone, D.J., Anderson, D., Shimkets, R.A., Burgess, C.E.,
Zerhusen, B.D., Liu, X., Spytek, K.A., Casman, S.J., Boldog, F.L.,
Smithson, G., Li, L. and Ji, W.
TITLE: Polypeptides and nucleic acids encoding same
JOURNAL: Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source
1..8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 4.52e-111 Length: 8354
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_2400_2600 (1-201) x AX556500 (1-8354)
QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7232 ATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCACATGGCGCGCGA 7291
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7292 GATTATGATGTGCTGGCCGACGCTGGACTAGCCACAGACCATGCTGTGGAAGCACCTT 7351
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7352 AGTAGCAGCAACGTGTCATGCCTTTTAATCTCTATATGTTCAAAAACAACAACCCCATCAGC 7411
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7412 AACTCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTTGA 7471
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7472 TTCCAGCTACACAACGCTGATCCCTGGTTATCCCAAACCCAGACATGGATGCCATGGAAACC 7531
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7532 TCCTACGAGCTCATCCACACACAGATGAAAACGCAGAGTGGGACAAACAGCAAGTCTATC 7591
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7592 CTCGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCTTTGTACCTTAGAACCGGTTT 7651
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 7652 GACCAGCTCTATGGCTCCACAATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAGTTT 7711

QY	161	AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal	180
Db	7712	GCATCCAGCGGCTCAGTCTTGGCAAGGGGTCAAGTTGCCTTGAAGGATGGCCGAGTG	7771
QY	181	ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn	200
Db	7772	ACCACAGACATCATCAGTGTGGCCAATGAGGATGGCGGAAGGGTTGCTGCCATCTTGAAC	7831
QY	201	His	201
Db	7832	CAT	7834
RESULT 4			
AB037723			
LOCUS	AB037723	8624 bp	mRNA linear PRI 10-MAY-2002
DEFINITION	Homo sapiens mRNA for KIAA1302 protein, partial cds.		
ACCESSION	AB037723		
VERSION	AB037723.2	GI:20521827	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	1 Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O. Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro		
AUTHORS	Nagase,T., Kikuno,R., Ishikawa,K.I., Hirose,M. and Ohara,O.		
TITLE	Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro		
JOURNAL	DNA Res. 7 (1), 65-73 (2000)		
MEDLINE	20181126		
PUBMED	10718198		
REFERENCE	2 (bases 1 to 8624)		
AUTHORS	Ohara,O., Nagase,T. and Kikuno,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute, Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba		
COMMENT	292-0812, Japan (E-mail:cdnaiffo@kazusa.or.jp, URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913, Fax:+81-438-52-3914)		
FEATURES	On May 9, 2002 this sequence version replaced gi:7242958.		
source	Location/Qualifiers		
gene	1..8624		
CDS	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="bf00104"		
	/tissue type="brain"		
	/clone_lib="pBCSKPGW"		
	/note="This sequence was replaced that of fg06864 cDNA as a representative cDNA sequence for KIAA1302."		
	1..8624		
	/gene="KIAA1302"		
	<590..3850		
	/gene="KIAA1302"		
	/note="Start codon is not identified."		
	/codon_start=1		
	/product="KIAA1302 protein"		
	/protein_id="BAA92540.2"		
	/db_xref="GI:20521828"		
	/translation="STERLDNKPETTRYDSFGRLTNVTFPTGOVSSFRSDTSSVHVQ VETSSKDDVTITNLSAGAFYLLQDQVRNSYIIGADGSLRLLLANGMEVALQTEPH LLAGTVPNTVGKRNVTLPIDNGLNLEWVRQKEQARGQVTFGRRLRVHNRNLLSLDF DRVTRTEKIYDDHRKFTLRILYDQAGPSLWSPSSRLNGVNVTVSPGGYIAGIQGIM SERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHSQRQYIFEFDKNDRLSSVTMP NVARQITLIRSVGYRNIYQPPGNASVIQDFTDGHLLHTFVLGTGRRVYIKYGL SKLAETLYDTKVSFTYDETAGMLKTNLQNEGFTCTIRYQIQLIDRQIFRFTEEG MVNARFDYNDNSFRVTSMQAVINETPLPIDLYRDYDVSQKTEQFKFGVIYDINQI ITTAVMTHTKHFDAYGRMKEVQYEIFSLMYWMTVQYDNMGRVVKELKVGPYANTTR YSYEYDADGQLQTVSINDKPLWRSYDNLNHLSPGNSARLTPURYDIRDRITRLG DVQYKMDGDFLRQGGDIFEYNSAGLLIKAYNRAGSWSVRVYDGLGRRVSSKSSHS HLQFFYADLTNPTKVTHLYNHSSEITSLYDLQGLFAMELSSGDEFYIACDNIQT PLAVFSGTGLMIKQILYAYGEIYMDTNPENFQIIGYHGGLYDPLTKLVHMGRRDYDV		

polyA_site
ORIGIN

Alignment Scores:
Pred. No.: 4.94e-111 Length: 8993
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x HSM806812 (1-8993)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 3088 ATAGGCTACCATGGTGGCCCTCTATGATCCACTCACCAAGCTTGTCACATGGGCCGCGA 3147

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 3148 GATTATGATGTGCTGGCCGACGCTGGACTAGCCACAGACACGAGCTGTGGAAGCACCTT 3207

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 3208 AGTAGCAGCAACGTCATGCCCTTTTAATCTCTATATGTTCAAAAACAACAACCCCATCAGC 3267

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 3268 AACTCCAGGACATCAAGTGTCTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGA 3327

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 3328 TTCCAGCTACACAACGTGATCCCTGGTTATCCCAAACCCAGACATGGATGCCATGGAACCC 3387

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 3388 TCCTACGAGCTTATCCACACACAGATGAAAACGCAGGAGTGGGACAAACAGCAAGTCTATC 3447

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 3448 CTCGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCTTTGTACCTTAGAACGGTTT 3507

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 3508 GACCAGCTCTATGGCTCCACAATCACCAAGTGCACGAGCTGCCAGAGGCTCCAAAGACCAAGAAGTTT 3567

QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 3568 GCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCTTGAAGGATGGCCGAGTG 3627

QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
Db 3628 ACCACAGACATCATCAGTGTGGCCAATGAGGATGGCGAAGGTTGCTGCCATCTTGAAC 3687

QY 201 His 201
Db 3688 CAT 3690

RESULT 7
AK056531
LOCUS
DEFINITION Homo sapiens cDNA FLJ31969 fis, clone NT2RP7008013, highly similar to Mus musculus mRNA for Ten-m4.
ACCESSION AK056531
VERSION AK056531.1 GI:16551956
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T., Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

Unpublished
2 (bases 1 to 2685)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source
Location/Qualifiers
1..2685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP7008013"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP7"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 5-weeks retinoic acid (RA) induction."
242..2434
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71206.1"
/db_xref="GI:16551957"
/translation="MLKTLNLQNEGFTCTIRYQIGPLIDRQIFRFTTEGMVNARFDYNYDNSFRVTSMQAVINETPLPIDLYRDDVSGKTEQFKFGVIYYDINQIITTAVMTKHFDAYGRMKVEQYEIIFRSLMYMTVQYDNMGRVVKELKVGYPYANTTRYSEYDAGQLQTVSINDKPLWRYSDLNGNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKXMDDGFLRQGGDIFEYNSAGLLIKAYNRAGWSVRYRIGLGRVSSKSSHHLQFFYADLTNPTKVTHLYNHSSEITSLYYDLQGLHFLAMELSSGDEFYIACDNIGTPLAVFSGTGLMKIQLYTAIGEIMYDTPNFQIIGYHGGLYDPLTKLVHMGRDDYDLAGRWTSPDHELWKHLSSSNVMPFNLYMFKNNNPISNSQDIKCFMTDVNSWLLTFGFLHNVI PGYPKPDMDAMEPSYELIHTQMKTEWDNSKSI LGVQCEVQKLFKAVTLERFDQLYGSTTSCIQAPKTKKFASSGSVFGKVFALKDGRVTDTIISVANEDGRVRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGLAILGLSGRRRTLENGVNVTVSQINTVLSGRTTRYTDIQLQYGALCLNTRYGTTLDEKARVLELARQRAVRQAWAREQQRLREGEGLRAWT
EGEKQQLSTGRVQGYDGFVISVEQYPELSDSANNIHFMRQSEMGRR"

ORIGIN
Alignment Scores:
Pred. No.: 7.2e-111 Length: 2685
Score: 1070.00 Matches: 200
Percent Similarity: 99.50% Conservative: 0
Best Local Similarity: 99.50% Mismatches: 1
Query Match: 99.35% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AK056531 (1-2685)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 1322 ATAGGCTACCATGGTGGCCCTCTATGATCCACTCACCAAGCTTGTCACATGGGCCGCGA 1381

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 1382 GATTATGATGTGCTGGCCGACGCTGGACTAGCCACAGACCAACGAGCTGTGGAAGCACCTT 1441

QY	41	SerSerSerAsnValMetPropPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer	60
Db	1442	AGTAGCAGCAACGTTCATGCCTTTTAATCTCTATATGTTCAAAAACAACCCCATCAGC	1501
QY	61	AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly	80
Db	1502	AACTCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTCACCTTTGGA	1561
QY	81	PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro	100
Db	1562	TTCCAGCTACACAACGTGATCCCTGGTTATCCCAACACGACATGGATGCCATGGAACCC	1621
QY	101	SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle	120
Db	1622	TCCTACGAGCTTATCCACACACAGATGAAAAACGCAGGAGTGGGACAACAGCAAGTCTATC	1681
QY	121	LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe	140
Db	1682	CTCGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCCCTTTGTACCTTAGAACGGTTT	1741
QY	141	AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe	160
Db	1742	GACCAGCTCTATGGCTCCACATCACCAGCTGCCTGCAGGCTCCAAAGACCAAGAGTTT	1801
QY	161	AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal	180
Db	1802	GCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTG	1861
QY	181	ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn	200
Db	1862	ACCACAGACATCATCAGTGTGGCCAATGAGGATGGCGAAGGTTGCTGCCATCTTGAAC	1921
QY	201	His	201
Db	1922	CAT	1924
RESULT 8			
AX675551			
LOCUS	AX675551	8438 bp	DNA linear PAT 27-MAR-2003
DEFINITION	Sequence 1 from Patent WO02055704.		
ACCESSION	AX675551		
VERSION	AX675551.1	GI:293333552	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1		
	Padigaru,M., Li,L., Zerhusen,B.D., Casman,S.J., Shenoy,S., Spytek,K.A., Zhong,M., Gangolli,E.A., Burgess,C.E., Patturajan,M., Vernet,C.A., Taylor,S., Tchernev,V.T., Miller,C.E., Guo,X., Boldog,F.L., Grosse,W.M., Alsobrook,J.P., Gerlach,V., Edingermark,S., Rothenberg,M.E., Ellerman,K., Macdougall,J., Malyankar,U., Millet,I., Peyman,J., Smithson,G., Gunther,E. and Stone,D.J.		
TITLE	Proteins, polynucleotides encoding them and methods of using the same		
JOURNAL	Patent: WO 02055704-A 1 18-JUL-2002; Curagen Corporation (US)		
FEATURES	Location/Qualifiers		
source	1..8438		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.76e-109	Length:	8438
Score:	1063.00	Matches:	201
Percent Similarity:	96.17%	Conservative:	0
Best Local Similarity:	96.17%	Mismatches:	0
Query Match:	98.70%	Indels:	8

DB:	6	Gaps:	1
US-10-029-020-14_COPY_2400_2600 (1-201) x AX675551 (1-8438)			
QY	1	IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg	20
Db	7259	ATAGGTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCACATGGCCGGCGA	7318
QY	21	AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu	40
Db	7319	GATTATGATGTGCTGGCCGACGCTGGACTAGCCAGACCACGAGCTGTGGAAGCACCTT	7378
QY	41	SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer	60
Db	7379	AGTAGCAGCAACGTTCATGCCTTTTAATCTCTATATGTTCAAAAACAACCCCATCAGC	7438
QY	61	AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly	80
Db	7439	AACITCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCACCTTTGGA	7498
QY	81	PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro	100
Db	7499	TTCCAGCTACACAACGTGATCCCTGGTTATCCCAACACGACATGGATGCCATGGAACCC	7558
QY	101	SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys	118
Db	7559	TCCTACGAGCTCATCCACACACAGATGAAACCGCAGGAGTGGACAACAGCAAGTAATT	7618
QY	119	-----SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLys	132
Db	7619	CCTGCACAAGGCTGCCAGTCTATCCTCGGGTACAGTGTGAAGTACAGAAGCAGCTCAAG	7678
QY	133	AlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGln	152
Db	7679	GCCTTTGTCACTTAGAACGGTTTGACCAGCTCTATGGCTCCACAATCACCAGTGCCAG	7738
QY	153	GlnAlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLys	172
Db	7739	CAGGCTCCAAAGACCAAGAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAG	7798
QY	173	PheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGly	192
Db	7799	TTTGCTTTGAAGGATGGCCGAGTGACCACAGACATCATCAGTGTGGCCAATGAGGATGGG	7858
QY	193	ArgArgValAlaAlaIleLeuAsnHis	201
Db	7859	CGAAGGTTGCTGCCATCTTGAACCAT	7885
RESULT 9			
AK122490			
LOCUS	AK122490	5583 bp	mRNA linear ROD 15-MAR-2003
DEFINITION	Mus musculus mRNA for mKIAA1302 protein.		
ACCESSION	AK122490		
VERSION	AK122490.1	GI:28972711	
KEYWORDS	FLI_CDNA.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1		
	Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Aizawa,H., Yuasa,S., Nakajima,D., Nagase,T., Ohara,O. and Koga,H.		
TITLE	Prediction of the coding sequences of mouse homologues of KIAA gene: II. The complete nucleotide sequences of 400 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries		
JOURNAL	DNA Res. 10, 35-48 (2003)		
REFERENCE	2 (bases 1 to 5583)		
AUTHORS	Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-FEB-2003) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan		

(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing: Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.

FEATURES

source

Location/Qualifiers
1. .5583
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg04788"
/tissue_type="brain"
/dev_stage="adult"
/note="vector:modified pBC SK+"

gene

1. .5583
/gene="mKIAA1302"

CDS

<1. .3597
/gene="mKIAA1302"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/codon_start=1
/evidence="not experimental"
/product="mKIAA1302 protein"
/protein_id="BAC65772.1"
/db_xref="GI:28972712"
/translation="TQNMVELSSPIDQELVLFDTSGKHLTYQSLPTGDLVYNFTYTGD
GDITHITDNNGNMNVRRDSTGMPLWLVPDQGVVWVTMGTSALRSVTTQGHELAMM
TYHGNSGLLATKSNENGWTFEYDPSFGLTNVTFPTQVSSFRSDTSSVHVQVETS
SKDDVTITTNLSAGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAG
TVNPTVGKRNVTLPIDNGLNVEWRQKEQARGQVTFGRRLRVHNRNLLSLDFDRVT
RTEKIYDDHKKFTIRILYDQGRPSLWSPSRLGNVNTVSPGGHIAGIQRGIMSERM
EYDQAGRITSRIADGKMSYTYLEKSMVLHLHSQRQYIFEDKNDRLSSVTMPNVAR
QTELTIRSVGYRNIYQPPEGNASVIOQDFTEDGHLHTFYLGTRRRVIYKYGKLSKLA
ETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCTIRYQIGPLIDRQIFRFTTEGMVNA
RFDYNDNSFRVTSQAVINETPLPIDLYRDDVSGKTEQFGKFGVIYYDINQIITTA
VMTHTKHFDAYGRMKVQVEIFRSLMYMTVQYDNMGRVVKELKVGPIYANTRYSYE
YDADGQLQTVSINDKPLWRYSYDLNGLNHLSPGNSARLTPLRYDLRDRITRLGDVQY
KMDEDFGLRQGGDFEYNSAGLLIKAYNRASGWSVRYDGLGRRVSSKSSHHLQ
FFYADLTNPTKVTHLYNHSSSEITSLYDLQGLHFLAMELSGDEFYIACDNICTPLAV
FSGTGLMIKOILYTAGYGIYDNTNPNFIIGYHGGLYDPLTKLVHMRDRDYDLAGR
WTSPPDEHLWKRLLSNSIVPFHLYMFKNNPNISNQDILKFMFTDVNSWLLTFGFOLHN
IPGYPKPDTDAMEPSYELVHTQMTQEWONSILGVQCEVQKQKAFVTLEREDQLY
GSTITSCQAPETKFKFASGSI FGKGVKFAKDGRTVTDIISVANEDGRRRIAILNNA
HYLENLHFTIDGVDTDFYFKPGPSEGDLAILGLSGRRLENGVNVTVSQINWLSGR
TRRYTDIQLQYRALCLNTRYGTTVDEKVRVLELARAVRQAWAREQQRLREGEGL
RAWTDGEKQQLVNTGRVQGYDGFVTSVEQYPELSDSANNIHFMRQSEMGR"

ORIGIN

Alignment Scores:

Pred. No.: 4.49e-106 Length: 5583
Score: 1031.00 Matches: 189
Percent Similarity: 99.00% Conservative: 10
Best Local Similarity: 94.03% Mismatches: 2
Query Match: 95.73% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AK122490 (1-5583)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
|||||
Db 2485 ATCGGCTACCAACGGCGCCTCTATGATCCACTCACCAGCTTGTCACATGGCGCCGACGG 2544
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
|||||
Db 2545 GATTATGATGTGCTGGCTGGACGCTGGACAGCCAGACCATGAACCTCTGGAACGCCTG 2604
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
|||||
Db 2605 AGTAGCAACAGCATCGTGCCTTTTCATCTCTACATGTTTAAAGAACAAACACCCCATCAGC 2664
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
|||||
Db 2665 AACTCTCAGGACATCAAGTGCTTCATGACAGATGTCAACAGCTGGCTCCTCACCTTTGGA 2724

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
|||||
Db 2725 TTCCAGCTGCACAACTGATACCTGGCTATCCCAAGCCAGACACAGATGCCATGGAACCC 2784
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
|||||
Db 2785 TCCTACGAGCTCGTACACACACAGATGAAAACTCAGGAATGGGACACACAGCAAGTCTATC 2844
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
|||||
Db 2845 CTCGGGGTACAGTGAAGTTTCAAGCAACTCAAGGCTTTCGTTACCTTAGAACGCTTT 2904
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
|||||
Db 2905 GACCAGCTCTACGGCTCGACCATCACCAGCTGCCAACAGGCCCTTGAGACAAAGAGTTT 2964
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
|||||
Db 2965 GCCTCCAGTGGTTCCATCTTTGGCAAGGGGTCAGTTTGCCTTGAAAGATGTCGAGTG 3024
QY 181 ThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
|||||
Db 3025 ACCACTGACATCATCAGTGTGGCCAATGAGGATGGCGGAGGATCGCAGCCATCTTGAAC 3084
QY 201 His 201
Db 3085 AAT 3087

RESULT 10

AB025413

LOCUS

DEFINITION

AB025413

ACCESSION

AB025413.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. .8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"

gene

1. .8585
/gene="ten-m4"
188. .8503
/gene="ten-m4"
/codon_start=1
/product="Ten-m4"
/protein_id="BAA77399.1"
/db_xref="GI:4760782"

CDS

1. .8585
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Balb/c"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="adult"

translation="MDVKERKPYRLTERRDAERRYTSSSADSEEGKGPQKSYSSSET
LKAYDQARLAYGSRVKDMVPOEAEEFCRTGNTFTLRELGLGEMTPHGLTYRTDIGL
PHCGYSMGASSDADLEADTVLSPVLRWGRSTRSGRSSCLSSRANSNLTLTDTTEHE

NTETDHPSSLQNHPRLRTPPPLPHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDH
SLSGEPPAGSAQEPTHQAQDNWLLNSNIPLETRNLGKQPFGLGTLDNLIEMDILSASRH
DGAYSDGHFLFKPGGTSPLFCTSPGYPLTSSVYSPPPRLPRQSPRQSPAPFNLKKPS
KYCNWKCAALSAILISATLIVLLAYFVAMHLFGLNWHLOPMEGQMOMYEITEDTASSW
PVPTDVSLYPSGGTGLTETPDRKGKGAAGKPSILFPEDSFIDSIEIDVGRASQKIPP
GTFWRISOVFIDHPVHLKFNVSGLKAAVLGIYGRKGLPPSHQTQDFVELLDGRRLLTQE
ARSLEGPQRQSRGVPVPSHETGFIQYLDLSGIWHLAFYNDGKESEVVSFLTIAIESVD
NCPNSQYNGDNCISGTHCFLGFLGPDGCRASCPVLCSGNGQYMKGRCLCHSGWKGA
CDVPTNQCIDVACSSHGTCIMGTICINPGYKGESCEEVDMCDPTCSSRGVCVGRGEC
SVGWGNTCETPRATCLDQCSGHGTFPLDPTGLCNCDPSTWTHDCSIEICAAADGGHGV
KEGGTCRCECDGWMGAACDQACHPRCAEHGTCRDGKCECSPGWNGEHCITIAHYLDRVV
KEGCPGLCNGNGRCTLDLNGWHCVQQLGWRGTGCDTSMETGCGDKDNDGDGLVDCMD
PDCCQLQFLCHVNPLCLGSPDPLDIIQETQAPVSQNLNPFYDRIKFLVGRDSTHSIPG
ENPFDDGHACVIRGQVMTSDGTPLVGVNISFINNPLFGYTIHQDGSFGLVTNGGISI
ILRFERAPFITQEHTLWLPWDRFFVMETIVMRHEENIEPSCDLSNEARPNPVVSPSL
TSFASSCAERGPVPEIQALQEEIVIAGCKWRLSYLSSRTPGYKSVLRISLTHPTIPF
NLMKVHLMVAVEGRLEFRKWFAAAPDLSYPIWDKTDVYNQKVFGESEAFVSVGYES
CPDLILWEKRTAVLQGYEIDASKLGGWSLDKHALNIQSGILHKNGENQFVSQPPV
IGSIMNGRRRSISCPSCNGLADGNKLLAPVALTCGDSGLSYGVDFNYIRRIFPSGNV
TNILEMRNKDFRSHSPAHKYIYLATDPMGAVFLSDTNSRRRVFKVSTTVVKDLVKN
EUVAGTGQCLPFDTRCGDGGKATEATLTNPRGTVTKFGLIYFVDGTMIRRVQDNG
IISTLLGNDLTSARPLSCDSVMEISQVRLEWPTDLAINPDMVSLVLDNNVVLQISE
NHQVRIVAGRPMHCQVPGIDHFLLSKVATHATLESATALAVSHNGVLVYIAETDEKKN
RIRQVTTSGEISLVAGAPSGCDCKNDANCDFSGDDGYAKDAKLNTPSSLAVCADGEL
YVADLGNIRIRFIRKNKPFNTQNMVELSSPIDQELYLFDTSGKHLYTQSLPTGDYLY
NFTYITGDGDIITHITDNNGMNVNRDSTGMPLWLVPDGVVYVMTGMTNSALRSVTTQ
GHELAMMTYHGNSGLLATKSNENGTTFVEYDSFGRLTNVTFPTGQVSSFRSDTSSV
HVQVETSSKDDVTIITNLSAGAFYTLLODQVRNSYIIGADGSLRLLLLANGMEVALQT
EPHLLAGTVNPTVGRNVTLPIDNGLNLVEWRQKEAQGVTVFGRRLRVHNRNLS
LDFDRVTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVTVSPGGHIAQIOR
GIMSERMEYDQAGRITSIRIFADGKMSYTYLEKSMVLHLHSQRQYIPEFDKNRLSSV
TMNVARQTLTETIRSVGYRNIYQPPEGNASVIOQDTEHGHLHTFYLGTGRRVIYKY
GKLSKLAETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCTIRYQIGPLIDRQIFRT
BEGMVNARFDYNDNSFRVTSQAVINETPLPIDLYRVDVSGKTEQFGKFGVIYDI
NQIITTAVMTHTKHFDAYGRNKEVQYEIFRSLMYMTVQYDNMGRVVKELKVGYPAN
TRYSYEYDADGOLQTVSINDKPLWRYSYDLNGLHLLSPGNSARLTPLRYDLDRIT
RLGDVQYKMDDEGFLRQRGDVFYNSAGLLIKAYNAGSWSVRYDYDGLRVRVSXS
SHSHHLQFFYADLTNPTKVTHLYNHSSSEITSLYYDLQGHLFAMELSSGDEFYIACDN
IGTFLAVFSGTGLMIKQILYTAIGEYIMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRD
YDVLAGRWTSPDHMLWKRLLSNSIVPFHLYMFKNPNPISNSQDIKCFMTDVNSWLLTF
GFQLHNVIPGYPKPDTDAMEPSYELVHTQMTQEWNSKILGVQCEVQKQKAFVTL
ERFDQLYGSTITSQCAPETKKFASSGSIFGKGVKALKDGRVTTDIIISVANEDGRRI
AAILNNAHYLENLHFTIDGVDTHTYFVKPGPSEGDILGLSGGRTLENGVNVTVSQI
NTMLSGTRTRYTDIQLQYRALCLNTRYGTTVDEKVRVLELARQRAVRAWAREQQRL
REGEEGLRAWTDGEKQQVILNTRGVQYDGFVTSVQYPELSDSANNIHFMFQSEMGR
R"

ORIGIN

Alignment Scores:
Pred. No.: 7.52e-106 Length: 8585
Score: 1031.00 Matches: 189
Percent Similarity: 99.00% Conservative: 10
Best Local Similarity: 94.03% Mismatches: 2
Query Match: 95.73% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AB025413 (1-8585)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7391 ATCGGCTACCA CGCGGCTCTATGATCCACTCACCAGCTTGTCACATGGCCGACGG 7450
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 7451 GATTATGATGTGCTGGTGACGCTGGACAAGCCAGACCATGAACCTCTGGAACGCTG 7510
QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 7511 AGTAGCAACAGCATCGTCCTTTTCATCTCTACATGTTTAAGAACAAACACCCCATCAGC 7570
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThrPheGly 80
Db 7571 AACTCTCAGGACATCAAGTGCTTCATGACAGATGTCAACAGCTGGCTCCTCACCTTTTGA 7630

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7631 TTCAGCTGCACAACGTTGATACCTGGCTATCCCAAGCCAGACACAGATGCCATGGAACCC 7690
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7691 TCCTACGAGCTCGTACACACACAGATGAAAACCTCAGGAATGGGACAAACAGCAAGTCTATC 7750
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7751 CTCGGGGTACAGTGTGAAGTTCAGAAAGCAACTCAAGGCTTTCGTTACCTTAGAACGCTTT 7810
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 7811 GACCAGCTCTACGGCTCGACCATCACAGCTGCCAACAGGCCCTTGAGACAAAGAGTTT 7870
QY 161 AlaSerSerGlySerValPheGlyLysGlyVallysPheAlaLeuLysAspGlyArgVal 180
Db 7871 GCCTCCAGTGGTTCATCTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGATGGTCGAGTG 7930
QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
Db 7931 ACCACTGACATCATCAGTGTGCCAATGAGATGGCGGAGGATCGCAGCATCTTGAAC 7990
QY 201 His 201
Db 7991 AAT 7993

RESULT 11
AF059485

LOCUS AF059485 9722 bp mRNA linear ROD 15-AUG-1998
DEFINITION Mus musculus DOC4 (Doc4) mRNA, complete cds.
ACCESSION AF059485
VERSION AF059485.1 GI:3170614
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 9722)
AUTHORS Wang,X.Z., Kuroda,M., Sok,J., Batchvarova,N., Kimmel,R., Chung,P.,
Zinsner,H. and Ron,D.
TITLE Identification of novel stress-induced genes downstream of chop
JOURNAL EMBO J. 17 (13), 3619-3630 (1998)
MEDLINE 98315054
PUBMED 9649432
REFERENCE 2 (bases 1 to 9722)
AUTHORS Wang,X.-Z. and Ron,D.
TITLE Direct Submission
JOURNAL Submitted (14-APR-1998) Skirball Institute, New York University
Medical Center, 550 First Ave., New York, NY 10016, USA
FEATURES
source
1. .9722
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/cell_line="NIH-3T3"
/note="tunicamycin-treated"
1. .9722
/gene="Doc4"
/note="downstream of CHOP 4"
583. .9060
/gene="Doc4"
/note="similar to Drosophila melanogaster tenm/odz and
human gamma-hergulin; type II transmembrane protein"
/codon_start=1
/product="DOC4"
/protein_id="AAC31807.1"
/db_xref="GI:3170615"
/translation="MDVKERKPYRSLTRRRDAERRYTSSADSEEGKGPQKSYSSSET
LKAYDQDARLAYGSRVKDMVPEAEFEFCRTGTGNTFLRELGLGEMTPPHGLTYRTDGL

gene

CDS

PHCYSGMGASSDADLEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLLTDTTEHE
NTEGAPLHCSSASSTPIEQSPSPPPANESQRRLLGNGVAQPTPDSDEEEFVFN
SFLVKGSASLGVAAANDHPSSLQNHPRLRTPPPLPHAHTPNQHHAAASINSLNRGNFT
PRSNPSAPTDHSLSGEPAGSADPETHAQDNWVLNSKIPVETRNLGKQFPFLGTWQDN
LIEMDIFASRRDGAYSDFHFFKPGCTSPLECTTSPTYGLTSTVYSPPPRPLPRST
FSRPAFNLKPSKYCNWKAALSAILISATVILLAYFVAMEHLFGLNHLQPMEGQM
MYEITEDTASSWPVPTDVSYPGGTGLTETDRKKGAAEGKPSLFPEDSFIDSGEI
DVGRRASQKIPPGTFWRQVFIHPVHLKFNVSLGKAALVIGYGRKGLPPSHTQLDFV
ELLDGRRLITQEARSLGEPQORQSPVPSPSHETGFIQYLDSDGWHLAFYNDGKESEV
VSFLTALIESVDNCPSPCYGNDICISGTCFLGFLGPCDGRASCPVLCSNGQYKMG
RCLCHSGWGAECVPTNQCIDVACSSHGTCIMGTICINPGYKGESCEEVDCMDPTCS
SRGVCVRGECHCSVGGWGTNCETPRATCLDQCSGHGTFPLDPTGLCNCDPFWTGHDCSI
EICAADCGHGVCGGTCTCEDGWMGAACDQACHPRAEHGTCRDGKCECTPGWNGE
HCTIEGCPGLCNGRCRTLNLNGWHVCQLGWRGTGCDTSMETGCGDKDNDGDLVD
CMDPDCLQLCHVNPLCLGSPDLIIQETQAPVSOQNLNSFYDRILKFLVGRDSTHS
IPGENPFDGGHACVIRGQVMTSDGTPLGVNISFINNPLFGYTIISRDQGSFDLVNCG
ISIIILFERAPEITQETHLWLWDRFFVFMETIWMRHEENEIPSRDLSNFPARNPVVSP
SPLTSPASSCAEKGPVPEIQALOEIIVAGCKMRLSYLSRTPGYKSVVRISLTHPT
IPFNLKVLHMAVEGRLEFRKWFAAAPDLSYFIWDKTDVYNQKVGLSEAFVSUYE
YESCPDLILWEKRTAVLQGYEIDASKLGWSLDKHALNIQSGILHKGNGENQFVSQQ
PPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDENYIRRIFFS
GNVTNILEMSPAHKYVYLATDPMGAVFLSDTNSRRVFKVSTTVVXDLVKNSEVVA
GTGDQCLPFDTRCGDGKATEATLTNPRGITVDKGLIYFVDGTMIRRVQDNGIIST
LLGNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNLSYLVDNNVVLQISENHQV
RIVAGRPMHCQVPGIDQFLLSKVAIHATLESATLAVSHNGVLYIAETDEKINRIRQ
VTSGEISLVAGAPSGCDKNDANCDGSGDDGAKDAKLNTPSSSLAVCADGELYVAD
LGNIRIRFIRKNKPLNTQMYELSSPIDQELYLFDTSKGKLYTQSLPTGYLYNFY
TGDGDIITHITDNNGMVNRDSTGMPLWLVPDGVYVWVTMTGNTSALRSVTTQGH
AMMTYHNGSGLLTKSNENGWTTTFEYDSEFGLTNVTFPTQGVSSERSDTSVHVQV
ETSSKDDVTITTHLSGSAFYTLLOQVRSYIIGADGSLRLLLLANGMEVALQTEPHL
LAGTVNPTVGRNVTLPIDNLNLVEWRQEQARGQVTFVGPRLRVHNRNLLSLDFD
RVTRTEKIYDDHRKFTLRILYDQAGRPSPFSPSSRLNGVNVITYSPGGHIAQIGDIMS
ERMEYDQAGRITSIRIFADGKMSYTYLEKSMVLHLHSORQYIFEDKNDRLSSVTMPN
VARQTLTETIRSVGYRNIYQPEGNASVIQDFTEDGHLHTFYLGTGRRVIYKYGLS
KLAETLYDTTKVSFTYDETAGMLKTVNLQNEGFTCTIRYQIGPLIDRQIFRFTEGM
VNARFDYNDNSFRVTSQVNETPLPIDLYRDDVSGKTEQFGKFGVIYDINQII
TTAVMTHSKHFDAYGRMKEVQYETFRSLMYMTVQYDNMGRVVKELKVGYPYANTRY
SVEYDADGLOQTVSINDKPLWRYSYDLNGLHLLSPGNSARLTPLRYDLRDRITRLGD
VQYKMEDEGSLRQGGDVFEYNSAGLLIKAYNRASGWSVRYDGLGRRVSSKSSH
HLQFFYADLTNPTKTVTHLYSHSSSEITSLYDQLQHLFAMELSGDEFFYIACDNIGTP
LAVFSGTGLMIKQILYATYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDYDVL
AGRWTSPDHELKRLSSNISVPHLYMFKNNPINSQDIKCFMTDVSNSWLLTFGFQL
HNVTGYPKPDTDAMEPSYELVHTQMTQEWDNKSILGVCEVQKQKAFVTLERFD
QLYGSITSCQAPETKKFASGSIKGVKFKALDGRVTTDIISVANEDGRRIAAIL
NNAHYLENLHFTIDGVDTYFVKPSPSGDLAILLGSRRRTLENGVNVTVSQTML
SGRTRTYDIIQYRALCLNRYGTVDEEKVRVLELQARQVQAWAREQQRLREGE
EGLRAWTDGEKQQVLNTRGVQYDGFVFTSVQYPELSDSANNIHFMRQSEMGRR"
1813. .1881
/gene="Doc4"
/note="putative; transmembrane-region site"

ORIGIN

Alignment Scores:
Pred. No.: 8.73e-106 Length: 9722
Score: 1031.00 Matches: 189
Percent Similarity: 99.00% Conservative: 10
Best Local Similarity: 94.03% Mismatches: 2
Query Match: 95.73% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x AF059485 (1-9722)

QY 1. IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
DB 7948 ATCGGCTACCACGGCGGCTCTATGATCCACTCACCAAGCTGTCCACATGGGCCGACGG 8007
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
DB 8008 GATTATGATGTGCTGGCTGGACGCTGGACAAAGCCAGACCATGAACCTCTGGAAACGGCTG 8067
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
DB 8068 AGTAGCAACAGCATCGTGCCTTTTTCATCTCTACATGTTTAAAGAACAAACACCCCATCAGC 8127

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
DB 8128 AACTCTCAGGACATCAAGTGTCTTATGACAGATGTCAACAGCTGGCTCCTCACCTTTGGA 8187
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
DB 8188 TTCCAGCTGCACAACTGATACCTGGCTATCCCAAGCCAGACACAGATGCGCATGGAACCC 8247
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
DB 8248 TCCTACGAGCTCGTACACACACAGATGAAAACTCAGGAATGGGACAAACAGCAAGTCTATC 8307
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
DB 8308 CTCGGGTACAGTGTGAAGTTCAGAACCACTCAAGGCTTTCGTTACCTTAGAACGCTTT 8367
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
DB 8368 GACCAGCTCTACGGCTCGACCATCACCAGCTGCCAACAGGCCCTGAGACAAAGAGTTT 8427
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
DB 8428 GCCTCCAGTGTTCATCTTGGCAAGGGGTCAAGTTTGCTTGAAGATGGTCGAGTG 8487
QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
DB 8488 ACCACTGACATCATCAGTGTGCCAATGAGGATGGCGGAGGATGCGAGCATCTTTGAAC 8547
QY 201 His 201
DB 8548 AAT 8550
RESULT 12
AB026980 9264 bp mRNA linear VRT 05-NOV-1999
LOCUS AB026980
DEFINITION Danio rerio mRNA for ten-m4, complete cds.
ACCESSION AB026980
VERSION AB026980.1 GI:5307784
KEYWORDS ten-m4.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (sites)
AUTHORS Mieda,M., Kikuchi,Y., Hirate,Y., Aoki,M. and Okamoto,H.
TITLE Compartmentalized expression of zebrafish ten-m3 and ten-m4,
homologues of the Drosophila ten(m)/odd Oz gene, in the central
nervous system
JOURNAL Mech. Dev. 87 (1-2), 223-227 (1999)
MEDLINE 99425191
PUBMED 10495292
REFERENCE 2 (bases 1 to 9264)
AUTHORS Mieda,M.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1999) Michihiro Mieda, Brain Science Institute,
RIKEN, Lab. for Developmental Gene Regulation; 2-1 Hirosawa,
Wako-shi, Saitama 351-0198, Japan (E-mail:mieda@brain.riken.go.jp,
Tel:81-48-467-9713, Fax:81-48-467-9714)
COMMENT Sequence updated (29-Jun-1999).
FEATURES Location/Qualifiers
source
1..9264
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/dev_stage="embryo"
1..9264
/gene="ten-m4"
323..8797
/gene="ten-m4"
/note="similar to Drosophila melanogaster tenm/odz and Mus
musculus Doc4"

```
/codon_start=1
/product="ten-m4"
/protein_id="BAA81893.1"
/db_xref="GI:5307785"
/translation="MEVKRRPYRSLTSSRRDTERRYTSSADSEDGKINPKSYSSSET
LKAFDQDRLAYGSRVKDLVHHEADFESRQPDFSLRDMAFGDPVPPHMGAYRTEMGL
PHRYSVSVASDADTETDGMSPHARVLWGRSNTKGRSSCLFSRANSNLTLTDEH
ENTENGPLHCSASSSPVDPSPPPSHAAQSQGRLLGNSAQAGRDSESEDEFPN
SFLVKTGSGNVCTPAAATANEQSFQNHSLRTPPLLFHSHSPSQHHTASIGLSRSN
YTQRNPSPAPTDSSAPNEGPTSAQSSAQNWLNSNVPLETRNIATQFLETLQD
NFIEMDILATARRDGAAYTDHFLFKPGGTSPLYCTTSPGYLTSSTVSPPPRPLRN
TFSRPAFLSKKPYKHCNWKCAALSAILISVTLVFLAYFIAMHLFGLNWHLPVQRQI
YQLTEDNTSGHLPTDLGLPPLGNTGLEFPDRGRDDGKLDGFFPEDSFIDMGEIDVG
RKVAQLIPPGIFWRSQVFIIDHMYLKNVLSKDALVGIYGRGLPPSHTQDFVELL
DGRLLSQGLPLDGPFPFAQORSLVPIITSHDTGCIQYMDSGIWHLAVYNDGKETEQV
SFLTAIDSIDDPCSNCFNGDCVSGNCHCFPGFRGPDCSRASCPLVCSGNGQYLKGR
CMCHSGWKGSECDVPTNQCIDITCSGHGTCTVGTICNPSYKGENCEEVDCLDPTCSG
RGVCRGECHECFVGGPGCESPRASCMEQCSHGFSFLADTNTCNDHNWTGHDGSTH
LCAADCGHGICVAGSCRCDEGMGTGCEQRACHPRCSHGTCCKDGKCECSPGWNGEH
CTIEGCPGLCNGNRCITLGNNGWYCVQLGWRGAGCDTSMETACSDGKNDGDGLTDC
MDPDCCLQASCHTTSLSVGSPPDLIDIIQETQISSSLSTLQSFYQRIHFLVGRDSTHVI
PDVNPFDGICHACVIRQVTSDDTPLVGVNISFINPKPAYGTYITRQDGSFDLVSNGGV
AIGLRFERAPFITQEHTLWLPWGRFFVMDTIIVMRHEVNDIPSCDLSFSTRPMPIVLP
PLTAFAGTCPERGIVPEIQTLQEEVRIPGTDMRLGLSSRTSGYKSLRLITLTHSTI
PESLMKVLHMAVEGRFLFRKWFSAAPNLSYDFVWDKTDVYSQVYGLSEAFVSUGFEY
ESCPDLILWEKRTAVLQYETTASNLGGSVDKHAALNIQSGILHKNGENIFTSQOP
PVGISIMGNRRRSISCPSCNGLADGNKLLAPVALACSGDGLSYGVDFNYVRRIFTTG
NVTSVLELSNPAHKYYLATSPVSGWLYLSDTSRKFVKVKSLYAVKDVAKNLELVAG
TGDQCLPYDETRCGDGGKAVEATLTNPRGITVDKYGVIFFDGTMIRRIDQNGHISTL
LGFNDLTSARPLSCDSVMDISQVRLEWPTDLAVSPMDNSLYLDNNVVLQISENHQVR
IVAGRPMHCQVPLDHLVSKIAIHATLESANALAVSHNGLLYIAESDEKIKNRVQV
STNGEISLLAGAPSGCDKNANDCDYSGDDGYAKDAKNAPSSLAVSPDGLFIADL
GNIRIRYVRNRKAFNLPLNMYEISSPIDDELXLFDVNASHVFTQSLTTGDYLYNFTYS
GEGDLSSITDKNKNRVSIRRDSTGLPLWLMGPDGQTFWFTGMTNNALKSVAAQGEIA
VMTYHGSGLLATKSNEDGWSTFYEYDNYGRLTNVITYPTGRVSSYRTSDSDTFRVQTE
GSNKEDITVTNLSASCTFYTLMDQOVKNYSYIIGDGLSLRLVLANGMEVSLHTEPHLL
SGTVNPTISKRNVTLPIDNGLNLVWRQRKEQAGQVTVYGRRLRVHNRNLLSMDFDR
VTRTEKYDDHRKFTLRIHYDHAGRPTLWAPSSRLNGVNVITYSPGGHIAQIQRGTMVS
RMEYDQNGRITSKIIFADGKSWSYTYLEKSMVLLYSQRYIFEFDKNRLSSVTMPNV
ARQLETTRSIGYVNTYRPPPEGNAVLQDYESDGLLLQTIHQGTGRRVITYKYGKLSR
LLEILYDTTRIAFSYDESAGMLKTVGLQSEGFACITRYRQIGPLIDRQIFRFSSEGMV
NARFDYNDNSFRVTSMQAVINETPLPIDLYRDDVSGKTEQFGKFGVIYDINQIIT
TAYMTHTKHFDAYGRVKEVQYEIFRSLMYWMVQFDMGRVVAKELVKVPYANTTRYA
YEYDADGQLQVVSINDKPLWYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDV
QYRLDEDEGFLRQNGDFFEYNSAGLLVKTYNKNGWTIKYRYDGLGRRVSSRSTQGH
LQFFYADLSSPTRVTHMYNHSSEITSLYDLQGHFLAMELSSGDEFYVACDNIGTPL
AVFSGAGMIKQILHTAFGEVYLDNSPFSQVLVIGYQGLYEPLTKLVHMGRRDYDVL
GRVTPPDHDIRKRLNSDNIYPNLYMFKNNPLSNSQETKCYMTDVSMLVTFEGFLY
NWIPGYRKPVTDAMEPSYELVHTQIKTQEWDSKSVLGQVCEVQRQLKSFVRLERFGQ
IYSASDSCGPTPLHTLFTATGTSFLFGKVKVAREGREVEADIISLANEDGRRRIAALV
KASYLQDLHETIAGLDTHYFVKSGLVEGDLSLLGMTVQRTLETGVNVTVSQVNMVLG
GRSRRITDIQMGYGLSLNVRYGSSVDEKRVLELARQAVATAWAHERHRLRQEE
GSRAWTDGERQQLLSSGRVQYEGFYIVSVDPFPELTNDINNHFWRQTEMGRR"
```

ORIGIN

Alignment Scores:			
Pred. No.:	1.87e-80	Length:	9264
Score:	807.00	Matches:	143
Percent Similarity:	87.50%	Conservative:	32
Best Local Similarity:	71.50%	Mismatches:	25
Query Match:	74.93%	Indels:	0
DB:	5	Gaps:	0

US-10-029-020-14_COPY_2400_2600 (1-201) x AB026980 (1-9264)

QY	1	IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg	20
Db	7685	ATAGGCTACGAGGCGGCCTTTATGAGCCCTTACAAAGCTTGTCATATGGCCGCAGA	7744
QY	21	AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu	40
Db	7745	GATTATGATGTTCTTGAGGCCCGGTGGACCACTCTGACCATGATATCCGGAAGCGGCTC	7804

QY	41	SerSerSerAsnValMetPropPheAsnLeuTyrMetPheLysAsnAsnProIleSer	60
Db	7805	AACAGCGACAAACATGTGCCCTTCAACCTCTACATGTTTAAGAACAAACCCCTTGAGC	7864
QY	61	AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly	80
Db	7865	AACAGCCCAAGAAACCAAGTGCTACATGACACAGATGTTAACAGCTGGCTGGTGACGTTCCGT	7924
QY	81	PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro	100
Db	7925	TTCCAGCTATACAACGTCATCCCTGGCTACCGCAAGCCTGTCAAGATGCCATGGAGCCT	7984
QY	101	SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle	120
Db	7985	TCGTACGAGTCGTTCCACACCCAGATAAAGACTCAGGAGTGGGATAGCACCAAGTCTGTT	8044
QY	121	LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe	140
Db	8045	CTAGGGGTGCAGTGTGAAGTCCAGAGCAGCTGAAGTCTTTTGTCCGGTTGGAGCGGTTT	8104
QY	141	AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe	160
Db	8105	GGTCAGATCTACAGTGCAAGTGATCTCGATGTCTCTCCAACCCCTCTCCACACACTTTT	8164
QY	161	AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal	180
Db	8165	GCAACAGGGACTTCTCTCTTCGGGAAAGGTGTGAAAGTCGCCATCCGTGAAGGCCCGGT	8224
QY	181	ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn	200
Db	8225	GAAGCGGACATCATCAGCCTGGCTAATGAAGATGGCCGAAGATCGCCGAGTCCTAGAC	8284

RESULT 13
GGA238613

LOCUS	GGA238613	8118 bp	mRNA	linear	VRT 02-JUN-1999
DEFINITION	Gallus gallus mRNA for teneurin-1.				
ACCESSION	AJ238613				
VERSION	AJ238613.1	GI:4877312			
KEYWORDS	ten-1 gene; teneurin-1.				
SOURCE	Gallus gallus (chicken)				
ORGANISM	Gallus gallus				
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.			
AUTHORS	Minet,A.D., Rubin,B.P., Tucker,R.P., Baumgartner,S. and Chiquet-Ehrismann,R.				
TITLE	Teneurin-1, a vertebrate homologue of the Drosophila pair-rule gene ten-m, is a neuronal protein with a novel type of heparin-binding domain				
JOURNAL	J. Cell. Sci. 112 (Pt 12), 2019-2032 (1999)				
MEDLINE	99276585				
PUBMED	10341219				
REFERENCE	2	(bases 1 to 8118)			
AUTHORS	Chiquet-Ehrismann,R.				
TITLE	Direct Submision				
JOURNAL	Submitted (28-APR-1999) Chiquet-Ehrismann R., R-1066 446, Friedrich Miescher-Institute, Postfach 2543, CH-4002 Basel, SWITZERLAND				
FEATURES	Location/Qualifiers				
source	1. .8118	/organism="Gallus gallus"			
		/mol_type="mRNA"			
		/db_xref="taxon:9031"			
gene	/tissue_type="brain, retina"				
	1. .8118	/dev_stage="E11, E14, adult"			
		/gene="ten-1"			
CDS	1. .8118	/gene="ten-1"			
		/function="putative role in neuronal development"			
		/codon_start=1			
		/product="teneurin-1"			
		/protein_id="CAB43098.1"			


```
repeat_unit      6667.. .6729
                  /gene="ten-1"
repeat_unit      6730.. .6801
                  /note="YD repeat 24"
repeat_unit      6802.. .6867
                  /gene="ten-1"
                  /note="YD repeat 25"
                  /gene="ten-1"
                  /note="YD repeat 26"
```

ORIGIN

```
Alignment Scores:
Pred. No.:      9.85e-65      Length:      8118
Score:          667.50      Matches:      118
Percent Similarity: 77.61%      Conservative: 38
Best Local Similarity: 58.71%      Mismatches: 44
Query Match:    61.98%      Indels:      1
DB:             5      Gaps:      1
```

US-10-029-020-14_COPY_2400_2600 (1-201) x GGA238613 (1-8118)

```
QY      1   IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db      7009 ATTGGTTTCATGGAGGCTCTATGATCTCTCACCAATTGGTTCATCTGGTCAAGCG 7068

QY      21   AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db      7069 GATTATGATGTTATTGCTGGTCGGACAAACACCAACCATCATATATGGAACACCTG 7128

QY      41   SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db      7129 AATGCTGTC---CCACAACCAATTCATCTACTCATTTGAAATAACTACCCAGTTGSC 7185

QY      61   AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db      7186 AGGATCCAAGATGTTGCTAAGTATACAACAGACATTGGGAAGTTGGCTAGAGCTATTGGT 7245

QY      81   PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db      7246 TTCCAGTTGCACAATGTTACTACCTGGGTTTCCAAAACCAAGAAATCGAAGCTTTGGAGACA 7305

QY      101  SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db      7306 ACATATGAACCTTCTACAGCTTCAACCAAAAACCCAGGAGTGGGATCCTGGAAAGACTATC 7365

QY      121  LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db      7366 CTTGGTATTTCAGTGTGAGTACAGAAGCAACTCCGAAACTTTATATCCTTGGATCAACTT 7425

QY      141  AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db      7426 CCAATGACCCCGGATAGTATAGTGCAAGTGTCTATGAGGGAGTGAAGCAACCGAGGTTT 7485

QY      161  AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db      7486 GCAGCTATTCTTCAGTATTGGAAGGCAATCAATTTGCTATCAAGGATGGCATCGTC 7545

QY      181  ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaLeuAsn 200
Db      7546 ACAGCGGACATTATTGGTGTGGCTAATGAGGACAGCGCGGCATCGCTGCCATACTCAAC 7605

QY      201  His 201
Db      7606 AAT 7608
```

```
RESULT 14
AB025410
LOCUS      AB025410
DEFINITION Mus musculus mRNA for Ten-m1, complete cds.
ACCESSION AB025410
VERSION   AB025410.1 GI:4760775
KEYWORDS  Ten-m1.
SOURCE    Mus musculus (house mouse)
```

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (sites)

REFERENCE

AUTHORS

Oohashi,T., Zhou,X., Feng,K., Richter,B., Moergelin,M., Perez,M.T., Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.

Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues

J. Cell Biol. (1999) In press

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho, Okayama, Okayama 700-8252, Japan

(E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128, Fax: +81-86-222-7768)

FEATURES

Source

Location/Qualifiers

1..8373

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Balb/c"

/db_xref="taxon:10090"

/dev_stage="adult"

/tissue_lib="brain"

1..8373

/gene="ten-m1"

96..8291

/gene="ten-m1"

/codon_start=1

/product="Ten-m1"

/protein_id="BAA77396.1"

/db_xref="GI:4760776"

/translation="MEQTDCKPYQPLSKVKHMDLAYTSSDSEDEGCRKPRQSFNSRE
TLHEYNOELRRNYSQSRKRDVEKSTQEIFCETPPTLCSGYHTDMHSVRHGYYQLE
MGSDVDTEGAASPDHALRMWIRGMKSEHSSCLSSRANSALSLTDTHERKSDGENG
FKFSPVCCDMEAPADSAQDMQSSPHNQFTFRPLPPPPPPHACTCARKPPPTVDSLQR
RSMSTRSQSPAAPAPPTSTQDSVHLHNSWVLSNIPLETRHFLFKHGSSSAIFSA
SQNYPLTSNTVYSPPLPRSTFSRPAFTFNKPYRCNWKCTALSATAITVTLALL
AYVIAVHLFGLTWQLQPVGQIYANGISNGNPGTESMDTTYSPIGGRVSDKSEKKVFK
GRAIDTGEVDICAQVMQTIPPGLFWRFQITIHPIYLKFNISLAKDSLGLIYGRNIP
PTHQTFQVLMGKLVKQDSKSDDIQHSRNLIILTSLETQETFIYMDQGPWYLAF
YNDGKMEQFVLTALTEIMDDCSTNCNGECISGHCHCFPGFLGPCDCARDSCPVLC
GGNGEYKHCVRNGWKGPCDVPEEQIDPTCFGHGTCIMGVICVPGYKGEICEE
EDCLDPMCSSHGICVKGECHCSTGWGVNCEPPLPIQEQCSGHGTFLLDTGVCSDP
KWTGSDCSTELCTMECGSHGVCGRGICQCEGEGWPTCEERSCHSCHAEGQCKDKGC
ECSPGWEGDHCTIAHYLDAVRDGCPLGFCNGRCTLDQNGWHCVQCQWSGTGCNIVM
EMLCGDNLNDGDLTDCVDPDCCQSNQVSPLCQGSPPDLDLIQQSQPLFSQHTSR
LFYDRIKFLIGKDSHVVPQDISFDSRRACVIRGQVAVDGTPLVGVNVFLHSDYD
FTISRQGSFDLVAIGGISVVLIFDRSPFLSEKRTLWLPNQFIVVEKVIQRIVADA
PSCDISNFI SPNPIVLPSPLTSGGSCPERGTIVPELQVVEEIPIPSSFVRLSYLSS
RTPGYKTLRIILLTHSTIPVGMIKVHLTVSVEGRLTKQWFPAAINLVYTFANWKTDIY
GQKVGLEALVSVGYEYEMCEPEFILWEQRTVVLOGFEMDASNLGWSLNKHHIFNQ
SGIIHKGNGENMFISQPPVIAITMGNGHQRSVACTNCGPAHNKLFAPVALASGPD
GSVYVGFNFVRRIFPSGNSVSILELRNRDTRHSTSPAHKYIYAMDPMSESLYLSDTN
TRKVYKLSLVETKOLSKNFEVVAGTDQCLPFDQSHCGDGGKASEASLNSPRGITVD
RHGFIYFVDGTMIRRIDENAVITTVIGSNGLTSTQPLSCDSGMDITQVRLEWPTDLAV
NPMDSLYVLDNNIVLQISENRRVRIIAGRPIHCQVPGIDHFLVSKVAIHSTLESARA
ISVSHSGLLFIAETDERKVNRIQOVTNTGEISIIAGAPTDCDCKIDPNCDCFSGDGGY
AKDAKMAPSSLAVSPDGLTVADLGNVRIRTIISNQAHNLMDNMLYEIASPADQELYQ
FTVNGTHLTMNLITRDYVYVNTFYNAEGDLGAIITSSNGNSVHIRRDAGGMPLWLVP
GQVYWLTISSNGVLKRVSAQGNALMTPYNGTGLLATKSNENGWTVYEDPEGLT
NATFPTGEVSSFHSDEKLTKVALLDTSNRENVLMSTNLATSTIYILKQENTQSTYRV
SPDGLRVTFASGMEINLSEPHILAGAVNPTLGKCI SLPGEHANLIEWRORKEQN
KGNVSAFERRLRAHNRNLLSIDFDMTRTKIYDDHRKFTLRILYDQTRPILWSPVS
RYNEVNITYSPSGLVTFIQRTWNEKMEYDQSGKIIISRTWADGKIWSYTYLEKSMML
LHSQRRYIFEXDQSDCLLSVTMPSMVRHSLQTMLSVGYRNIYTPDSSTSFIDYSR
DGRLLQTLHLGTRRRVLYKTKOARLSEILYDQTVTLTYEESGVIKTIHLMDHGF
CTIRYRQTGPIGRQIFRFESEGLVNARFDYSNNFRVTSMQAVINETPLPIDLRYV
DVSGRTEQFGKFSVINYDLNQVITTTVMKHTKIFNANGVIEVQYEILKAIAYWMTIQ
YDNMGRMVICDIRGVDMANITRYFYEYDAGQLQTVSNDKIQWRYSYDLNGNLLIS
HGNSARLTPLRYDLRDRITRLGEIQYKMDDEGFLRQRGNDIFEYNSNGLLQKAYNKVS
GWTVQYYYDGLGRRRVASKSSLGHLQFFYADLANPIRVTHLYNHTSAEITSLYYDLQG

HLIAMELSSGEEYVACDNMGTPLAVFSSRGQVIKELLYTPYGDYVHDYTPDFEVIIG
FHGGLYDLTKLVHLGQRDYDVVAGRWTPNHHIWKQLNLLPKPFLNLYSFENNYPVGK
IQDVAKYTTDICTWLEFGFQLHNVLPGFPKPELENMELTYELLQOTKTKQEWDPGKM
ILGIQCEQLQQLRNFISLDQLPMTPOYNEGRCLGEGKQPRFAAVPSVFGKGIKFAIKE
GIVTADIIIGVANEDSRRLAAILNNAHYLENLHFTIEGRDTHFYIKLGSLEEDLVLIGN
TGGRRILEGNVNTVSQMTSVLNGRTRFADIQLOHGALCFNIRYGTTVEEERKNHVLE
MARQRAVAQAQWTOEQRRLOEGEGEGRVWTEGEKQQLLGTGRVQGDGYFVLSVEQYLE
LSDSANNIHFMRQSEIGRR"

ORIGIN

Alignment Scores: 3.3e-61 Length: 8373
Pred. No.: 636.50 Matches: 116
Score: 75.37% Conservative: 37
Percent Similarity: 57.14% Mismatches: 45
Best Local Similarity: 59.10% Indels: 5
Query Match: 10 Gaps: 2
DB:

US-10-029-020-14_COPY_2400_2600 (1-201) x AB025410 (1-8373)

Qy	1	IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg	20
Db	7182	ARTGGTTTTCATGGAGGACTCTATGATTTTCTTACTAACTAGTCATCTAGGGCAAGA	7241
Qy	21	AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu	40
Db	7242	GATTATGATGTTGTTGCTGGCAGATGGACAACGCCTAATCATCATATGGAACAGTTG	7301
Qy	41	SerSerSerAsnValMet-----PropheAsnLeuTyrMetPheLysAsnAsnPro	58
Db	7302	-----AACCTTCTTCTTAAACCATTCACCTCTACTCTTTGAGAATAACTACCCA	7352
Qy	59	IleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThr	78
Db	7353	GTTGGCAAAATTCAGATGTTTGCAGAAAGTATACCACAGATATCGGCACCTGGCTGGAGCTC	7412
Qy	79	PheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet	98
Db	7413	TTTGGTTTCCAGTTACATAACCTACTACCTGGATTTCCCAACACAGAACTTGAAACATG	7472
Qy	99	GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys	118
Db	7473	GAATTAACCTATGAGCTTCTTCAACTCCAGACAAAACCCAGAGTGGGATCCTGGAAAG	7532
Qy	119	SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu	138
Db	7533	ATGATCCTGGGCATTCACTGTGAGCTCCAGAAACAACTAAGGAATTCATTCTCTTGAT	7592
Qy	139	ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys	158
Db	7593	CAGCTTCCTATGACTCCTCAGTACAATGAGGGGAGGTGTCTTGAAGGAGGGAACAGCCG	7652
Qy	159	LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly	178
Db	7653	AGGTTTGCTGCTGCTCCCTTCTGTCTTTGGCAAAGGTATCAAATTTGCCATCAAGGAGGC	7712
Qy	179	ArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIle	198
Db	7713	ATAGTACAGCTGATATTATAGGAGTAGCCCAATGAAGATAGCAGGCGTCTTGCTGCCATT	7772
Qy	199	LeuAsnHis	201
Db	7773	CTCAACAAT	7781

RESULT 15
AF100772
LOCUS AF100772 8297 bp mRNA linear PRI 30-NOV-1999
DEFINITION Homo sapiens tenascin-M1 (TNM1) mRNA, complete cds.
ACCESSION AF100772
VERSION AF100772.1 GI:6165844
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

TITLE

JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES

source

gene

CDS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 8297)
Brandau,O., Schuster,V., Weiss,M., Hellebrand,H., Fink,F.M., Kreczy,A., Friedrich,W., Strahm,B., Niemeyer,C., Belohradsky,B.H. and Meindl,A.
Epstein-Barr virus-negative boys with non-Hodgkin lymphoma are mutated in the SH2D1A gene, as are patients with X-linked lymphoproliferative disease (XLP)
Hum. Mol. Genet. 8 (13), 2407-2413 (1999)
20025749
10556288
2 (bases 1 to 8297)
Brandau,O., Ohashi,T., Faessler,R. and Meindl,A.
Direct Submission
Submitted (22-OCT-1998) Medizinische Genetik, LMU-Muenchen, Goethestr. 29, Muenchen 80336, Germany
Location/Qualifiers
1..8297
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xq24-q25"
1..8297
/gene="TNM1"
65..8242
/gene="TNM1"
/function="putative receptor molecule"
/codon_start=1
/product="tenascin-M1"
/protein_id="AAF04723.1"
/db_xref="GI:6165845"
/translation="MEQTDCKPYQPLPKVKHEMDLAYTSSSDESDGRKPKRQSYNSRE TLHEYNOELRMVNSQSRKRKEVEKSTQEMEFCTSHTLCSGYQYTMHVSVRHGYLE MGSVDVDTETEGAAAPDHALRMWIRGMKSEHSSCLSSRANSALSLTDTHERKSDGENG FKESPVCCDMEAQGSTQDVQSSPHNQFTFRPLPPPPPPHACTCARKPPPAADSLQR RSMTTRSQSPAAPAPTSTQDSVHLHNSWVLSNIPLRHSLEFKHSGSSAIFSAA SQNYPLTSNTVYSPPPRPLRSTFSRPAFTFNKPYRCNWKCTALSATAITVTLALIL AVVIAVHLFGLTWQLQPVGELYANGVSKGNRSMDDTYSPIGGKVDKSEKKVFQ KGRAIDTGEVDIGAQMOTIPPLFWRFQITIHPIYLKFNISLAKDLSLGIYGRNI PPHTQDFDFVKLMDGKQLVKQDSKGSDDTQHSRNLIITSLQETGFIEMDQGPWYLA FYNDGKMEQVFLVLTAAIEIMDDCSTNCNGEGECISGHCHCFPGFLGPDSCAPVL CGNGEYKGCVCVRHGWKGPEDVPEEQCIDPTCFGHGTCIMGVCICVPGYKGEICE BEDCLDPMCSNHGICVKGECSTCWGGVNCETPLPVCQECCSGHGTFLLDAGVCSCD PKWTGSDCSTELCTMECSHGVCVSRGICQCEGWVGPTEERSCHSCHTEHGCKDGK CECSPGWEGDHCTIAHYLDVARDGCPGLCFNGRCRTLDQNGWHCVQCVGSGTGNVV MEMLCGNDLNDGDLTDCVDPDCCQSNICYISPLCQSPDPLDLIQQSQTLSQHTS RLFYDRIKFLIGKDSHTVIPPVSVFDSRRACVIRGVVAIDGTPLVGNVSVFLHSDY GFTISRQDGSFDLVAIGGISVILIFDRSPFLPEKRTIWLPNQFIVVEKVTMQRVVSD PPSCDISNFI SPNPIVLPSPLTSGGGCPEGTIVPELQVVEEIPIPSSFVRLSYLS SRTPGYKTLRLILLTHSTIPVGMIKVHLTVAVEGRLTKQFPAAINLVYTFAMNKTDI YGQVWGLAEALVSQVEYETCPDFILWEQRTVVLQGFEMDASNLGDSWLNKHHILNP QSGIHKNGENMFISQPPVISTIMNGHQRSVACTNCNGPAHNNKLFAPVALASGP DGSVYVGFDFVRRIFPSGNSVSIELSTSPAHHYYLAMDPVSESLYLSDTNTRKVKYK LKSLVETKDLSKNFVAVGTGQCLPFDQSHCGDGRASEASLNSPRGITVDRHGFIY FVDGTMIRKIDENAVITTVIGSNGLTSTQPLSCDSGMDITQVRLIEWPTDLAVNPMDNS LYVLDNNIVLQISENRRVRIIAGRPVHCQVPGIDHFLVSKVAIHTLESARAI SVSHS GLLFIAETDERKNRIQQVTTNGEIIYIAGAPTDCKIDPNDCDFSGDGGYAKOAKM KAPSSIAVSPDGLTYVADLGNVRIRTSRNQAHLDNMIYIEIASPADQELLYQFTVNGT HLHNLNITRDVYNYFTYNSEGDLGAI TSSNGNSVHIRRDAGMPFLVVPGGQVYWL TISSNGVLKRVSAQGNPALMTYPGNTGLLATKSNENGWTTVVEYDPEGHLTNATFTT GEVSSFHSDLEKLTKEVLDTSNENVLMSTNLSTATSTIYLKQENTQSTYRNVNPDGSL RVTRFAGMEIGLSSEPHILAGAVNPTLGKCNISLPGEHANLIEWRQRKEQNKNVSA FERRLAHNRNLLSIDFDHITRTKIIDHRKFTLRILYDQTRPILWSPVSRVNEVNI ITYSPSGLVTFIQRGTWNEKMEYDQSGKII SRTWADGKIWSYTYLEKSVMLLLHSQRR YIFEYDQSDCLLSVTMPMSVRHSLQTMLSVGYYRNIYTPPDSSTSFIDYSRDRGLLQ TLHLGTGRVLYKYTKQARLSEVLYDTQVTLTYEESGVIKTHLMHDGFICTIRYR QTGPLIGRQIFRFSEEGLVNARFDYSYNFRVTSMQAVINETPLDILYRYVDVSGRT EQFGKFSVINYDLNQVITTTVMKHTKIFPSANGQVIEQVIEILKAAYMTIQYDNVGR HGNMCI RVGDANITRYFEYDADGQLQTVSVNDKTQWRSYDLNGDINLLSHGKSAR LTPRLYDLRDRITRLGEIQYKMEDEGFLRQGNDFEYNSNGLLQKAYNKASGWTQVY YYDGLGRRRVASKSSLGHLQFFVDATEPNFIRVTHLYNHTSSEITSLYDLQGHLIAME

LSSGEYYVACDNTGTPLAVFSSRGQVIKEILYTPYGDIIYHDTYPDQVILGFHGGLY
DFLTKLVHLGQDYDVVAGRWTTAYHHIWKQLNLLPKPFNLYSFENNPVGVKIQDVAK
YTDIRSWLELFGQLHNVLPGFPKPELENLELTYLQTKTQWDPGKTIILGIQC
ELQKQLRNFISLDQLPMTPRYNDGRCLGGKQPRFAAVPSVFGKIKFAIKDGIVTAD
IIGVANEDSRRLAAILNNAHYLENLHFTIEGRDTHYFIKLGSLIEDLVLIIGNTGGRRI
LENGVNVTVSQMTSLNNGRTRRFADIQLQHGAALCFNIRYGTVEEKNHVLLEIARQRA
VAQAWTKEQRRLQEGEGIRAWTEGEKQQLLSTGRVQGYDGYFVLVVEQYIELELSDSAN
NIHFMRQSEIGRR"

ORIGIN

Alignment Scores:
Pred. No.: 1.56e-60 Length: 8297
Score: 630.50 Matches: 116
Percent Similarity: 74.88% Conservative: 36
Best Local Similarity: 57.14% Mismatches: 46
Query Match: 58.54% Indels: 5
DB: 9 Gaps: 2

US-10-029-020-14_COPY_2400_2600 (1-201) x AF100772 (1-8297)

QY	1	IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg	20
Db	7133	ATTGGTTTTCATGGAGGACTCTATGATTTCCTTACTAAATTAGTGCACCTGGGCAAAGG	7192
QY	21	AspTyrAspValLeuAlaGlyArgTTrpThrSerProAspHisGluLeuTrpLysHisLeu	40
Db	7193	GATTATGATGTTGTGTGGCAGATGGACAACGGCCTATCATCATATATGGAACAGTTG	7252
QY	41	SerSerSerAsnValMet-----ProPheAsnLeuTyrMetPheLysAsnAsnPro	58
Db	7253	-----AACCTCCTTCCTAAACCATTCACCTCTACTCCTTTGAAAATAACTACCCA	7303
QY	59	IleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThr	78
Db	7304	GTTGGCAAAATTCAAGATGTTGCAAAAGTATACCACAGACATCAGAAGTTGGTTGGAGCTA	7363
QY	79	PheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet	98
Db	7364	TTTGGTTTCCAATTACACAATGTACTACCTGGATTTCCTCAAACTGAATTAGAAAAATTA	7423
QY	99	GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys	118
Db	7424	GAATTAACTTACGAGCTTCTACGGCTTCAGACAAAACCTCAAGAGTGGATCCTGGAAAG	7483
QY	119	SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu	138
Db	7484	ACTATCCTGGGCATTTCAGTGTGAACCTCCAGAAACAGCTCAGGAATTTTCATTTCCCTGGAC	7543
QY	139	ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys	158
Db	7544	CAACTACCTATGACTCCCGGATACAAATGATGGACGGTGCCTTGAAGGAGGGAAGCAACCA	7603
QY	159	LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly	178
Db	7604	AGGTTTGCTGCTGTCCCTTCTGTTTTTGGGAAAGGTATAAAATTTGCCATCAAGGATGGC	7663
QY	179	ArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIle	198
Db	7664	ATAGTAACAGCTGATATTATAGGAGTAGCCAAATGAAGATAGCAGGCGGCTTGCTGCCATT	7723
QY	199	LeuAsnHis	201
Db	7724	CTCAATAAT	7732

Search completed: August 14, 2004, 12:03:02
Job time : 3600.05 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 23:35:57 ; Search time 1344.2 Seconds
(without alignments)
2450.585 Million cell updates/sec

Title: US-10-029-020-14_COPY_2650_2725
Perfect score: 391
Sequence: 1 SQINTVLNGRTRYTDIQLQ.....EQQLREGEEGLRAWTEGK 76

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29275/app_query.fasta_1.3519
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029020@cgn 1 1 19065 @runat_06082004_112216_29275 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	391	100.0	3320	9	AK127101	AK127101 Homo sapi
2	391	100.0	3486	9	HSM806114	BX537983 Homo sapi
3	391	100.0	5269	9	HSM800629	AL080120 Homo sapi
4	391	100.0	8354	6	AX556500	AX556500 Sequence
5	391	100.0	8438	6	AX675551	AX675551 Sequence
6	391	100.0	8624	9	AB037723	AB037723 Homo sapi
7	391	100.0	8645	6	AX600210	AX600210 Sequence
8	391	100.0	8993	9	HSM806812	BX640737 Homo sapi
9	391	100.0	68970	2	AC090378	AC090378 Homo sapi
10	391	100.0	170156	2	AP002412	AP002412 Homo sapi
11	391	100.0	186084	9	AP002768	AP002768 Homo sapi
12	386	98.7	2685	9	AK056531	AK056531 Homo sapi
13	371	94.9	834	5	GGA293019	AJ293019 Gallus ga
14	368	94.1	208760	2	AC074048	AC074048 Mus muscu
15	368	94.1	226256	10	AC073599	AC073599 Mus muscu
16	365	93.4	5583	10	AK122490	AK122490 Mus muscu
17	365	93.4	8585	10	AB025413	AB025413 Mus muscu
18	365	93.4	9722	10	AF059485	AF059485 Mus muscu
19	364	93.1	231970	2	AC133064	AC133064 Rattus no
20	364	93.1	257693	2	AC103323	AC103323 Rattus no
21	359	91.8	68970	2	AC090378	AC090378 Homo sapi
22	311	79.5	134245	9	HS1052M9	AL022718 Human DNA
23	308	78.8	8297	9	AF100772	AF100772 Homo sapi
24	306	78.3	8118	5	GGA238613	AJ238613 Gallus ga
25	304	77.7	8816	5	AB026979	AB026979 Danio rer
26	304	77.7	188655	2	BX324137	BX324137 Danio rer
27	303	77.5	8373	10	AB025410	AB025410 Mus muscu
28	303	77.5	134334	10	AL163512	AL163512 Mouse DNA
29	303	77.5	165860	10	AL831716	AL831716 Mouse DNA
30	303	77.5	172465	2	BX324192	BX324192 Mus muscu
31	303	77.5	250093	2	AC114058	AC114058 Rattus no
32	296	75.7	2157	6	AX876360	AX876360 Sequence
33	296	75.7	2157	6	BD156088	BD156088 Primer fo
34	296	75.7	2157	9	AK027473	AK027473 Homo sapi
35	296	75.7	3270	6	AX877449	AX877449 Sequence
36	296	75.7	3270	6	BD156663	BD156663 Primer fo
37	296	75.7	3270	9	AK001748	AK001748 Homo sapi
38	296	75.7	3614	6	AX876525	AX876525 Sequence
39	296	75.7	3614	6	BD156175	BD156175 Primer fo
40	296	75.7	3614	9	AK001336	AK001336 Homo sapi
41	296	75.7	5804	10	AK122513	AK122513 Mus muscu
42	296	75.7	7816	10	AF195418	AF195418 Mus muscu
43	296	75.7	8473	6	AX662357	AX662357 Sequence
44	296	75.7	8487	6	AX662359	AX662359 Sequence
45	296	75.7	8645	6	AX662355	AX662355 Sequence

ALIGNMENTS

AK127101	AK127101	3320 bp	mRNA	linear	PRI 09-SEP-2003
LOCUS	Homo sapiens cDNA FLJ45158 fis, clone BRAWH3043034, highly similar to Mus musculus neuregulin 1 (Nrg1).				
DEFINITION					
ACCESSION	AK127101				
VERSION	AK127101.1	GI:34533862			
KEYWORDS	oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.				
TITLE	NEDO human cDNA sequencing project				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 3320)				
AUTHORS	Isogai,T. and Yamamoto,J.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)				
COMMENT	NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.				
FEATURES	Location/Qualifiers				
source	1..3320				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/clone="BRAWH3043034"				
	/tissue_type="brain"				
	/clone_lib="BRAWH3"				
	/note="cloning vector: pME18SFL3"				
ORIGIN					
Alignment Scores:	2.75e-33	Length:	3320		
Pred. No.:	391.00	Matches:	76		
Score:		Conservative:	0		
Percent Similarity:	100.00%	Mismatches:	0		
Best Local Similarity:	100.00%	Indels:	0		
Query Match:	100.00%	Gaps:	0		
DB:	9				
US-10-029-020-14_COPY_2650_2725 (1-76) x AK127101 (1-3320)					
QY	1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20				
Db	2706 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAG 2765				
QY	21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40				
Db	2766 TACGGGCACTGTGCTTGAACACACACGCTACGGGACAAACGTTGGATGAGGAGGACGCG 2825				
QY	41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrrAlaArgGluGlnArg 60				
Db	2826 GTCCTGGAGCTGGCCCGCAGAGAGCGCTGCCCAAGCGTGGGCCCGCAGCAGAGA 2885				
QY	61 LeuArgGluGlyGluGluGlyLeuArgAlaTrrThrGluGlyGluLys 76				
Db	2886 CTGCGGAGAGGGAGGAGGAGGCTGCGGCGCTGGACAGAGGGGGAGAG 2933				
RESULT 2					

HSM806114	HSM806114	3486 bp	mRNA	linear	PRI 17-JUN-2003
LOCUS	Homo sapiens mRNA; cDNA DKFZp686D0412 (from clone DKFZp686D0412).				
DEFINITION					
ACCESSION	BX537983				
VERSION	BX537983.1	GI:31874053			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS	1 (bases 1 to 3486) Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.				
TITLE	Direct Submission				
JOURNAL	Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY				
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp686D0412) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.				
FEATURES	Location/Qualifiers				
source	1..3486				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/map="11q14.1"				
	/clone="DKFZp686D0412"				
	/tissue_type="human cervix"				
	/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"				
	/dev_stage="adult"				
gene	1..3486				
	/gene="DKFZp686D0412"				
CDS	<1..3217				
	/gene="DKFZp686D0412"				
	/note="hypothetical protein, N-terminus truncated"				
	/codon_start=2				
	/product="hypothetical protein"				
	/protein_id="CAD97943.1"				
	/db_xref="GI:31874054"				
	/translation="SFGRLTNVTFPTGOVSSFSRSDTSSVHVQVETSSKDDVTITNL SASGAFYTLQDQVRNSYIIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVWRQKEQARGQVTFGRRLRVNRLNLSLGFDRVTRTEKIYDDHRKFTLRILYDOAGPSLWSPSSRLNGVNVYSPGGIAGIQRGIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLLHSQRYIFEFDKNDRLSSVTMPNVAROTLETIRSVGYRNIYQPPENASVIOQDFTEDGHLHFTYLGTRRRVIYKYGKLSKLAETFYDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDRQIFRTEEGMVNARFDYNDNSFRVTSMAVINETPLPIDLYRYDDVSGKTEQCFKFGVIYDINQIITTAVMTHTKHFDAYGRMKEVQYEIFRSMLYWMVTQYDNMGRVVKELKVGYPYANTTRYSYEYDADQLQTVSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEDGFLRQRGDDIFEYNSAGLLIKAYNRAGSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPTKVTHLYNHSSSEITSLYYDLQGHFLFAMELSSGDFYIACDNIGTPLAVFSGTGLMIKQILYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTKLVHMRDYYDLVLAGRTWSPDHELWKHLSSSNVMPFNLNFKNNNPISNSQDIKCFMTDVNSWLLTFGFQLHNVI PGYPKPDMDAMEPSYELIHTQMTQEDNDSKSLGVQCEVQKQKAFVTLERFDQLYSTITSQQAPKTKKFASSGSGVFGKVKFALKDGRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFKPGSPGDLAILGLSGGRRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDEEKARVLELARQRAVRQAWAREQQRLRVEEGEGLRAWTEGKQQVLSLGRVQGYDGGFFVISVEQPELSDSANNIHFMRQSEMGRR"				
	3468				
	/gene="DKFZp686D0412"				
polyA_site					
ORIGIN					
Alignment Scores:	2.9e-33	Length:	3486		
Pred. No.:	391.00	Matches:	76		
Score:					

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x HSM806114 (1-3486)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||||
Db 2855 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 2914
|||||

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
|||||
Db 2915 TACGGGGCACTGTGCTTGAACACACACGCTACGGGACAACTGGATGAGGAGGACACGG 2974
|||||

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
|||||
Db 2975 GTCCTGGAGCTGGCCCGCAGAGAGCCGTGCGCAAGCGTGGCCCGCGAGCAGACAGAGA 3034
|||||

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
|||||
Db 3035 CTGCGGGAAGGGAGGAAGGCCTGCGGGCCTGGACAGAGGGGGAGAAG 3082
|||||

RESULT 3
HSM800629 HSM800629 5269 bp mRNA linear PRI 18-FEB-2000
LOCUS Homo sapiens mRNA; cDNA DKFZp564O0423 (from clone DKFZp564O0423);
DEFINITION partial cds.
ACCESSION AL080120
VERSION AL080120.1 GI:5262552
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5269)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (16-JUN-1999) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp564O0423) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES
source Location/Qualifiers
1..5269
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp564O0423"
/tissue_type="brain"
/clone_lib="564 (synonym: hfbr2). Vector pAMP1; host
X1-2blue; sites NotI + Sali"
/dev_stage="fetal"
1..5269
/gene="DKFZp564O0423"
<1..556
/gene="DKFZp564O0423"
/note="strong similarity to C-terminus of mouse DOC4"
/codon_start=2
/product="hypothetical protein"
/protein_id="CAB45719.1"
/db_xref="GI:5262553"
/db_xref="SPTREMBL:Q9Y4S2"
/translation="GANEDGRRVAAILDHAHVLENLHFTIDGVDTHYFVKPSPSEGL
AIIGLSGRRITLGVNVTVSQINTVNGRTRRYTDIQLQYGALCLNTRYGTTLDEEK
ARVLELRQRAVRQAWAREQQRLREGEEGLRAWTEGEKQQLVSTGRVQYDGFVISV
EQYPELSDSANNIHFMRQSEMGRR"

gene
CDS

polyA_signal 5226..5231
/gene="DKFZp564O0423"
5251
polyA_site
/gene="DKFZp564O0423"

ORIGIN

Alignment Scores:
Pred. No.: 4.55e-33 Length: 5269
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x HSM800629 (1-5269)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||||
Db 194 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 253
|||||

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
|||||
Db 254 TACGGGGCACTGTGCTTGAACACACACGCTACGGGACAACTGGATGAGGAGGACACGG 313
|||||

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
|||||
Db 314 GTCCTGGAGCTGGCCCGCAGAGAGCCGTGCGCAAGCGTGGCCCGCGAGCAGACAGA 373
|||||

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
|||||
Db 374 CTGCGGGAAGGGAGGAAGGCCTGCGGGCCTGGACAGAGGGGGAGAAG 421
|||||

RESULT 4
AX556500 AX556500 8354 bp DNA linear PAT 27-NOV-2002
LOCUS Sequence 13 from Patent WO02057453.
DEFINITION AX556500
ACCESSION AX556500
VERSION AX556500.1 GI:25899736
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gangolli,E.A., Patturajan,M., Vernet,C.A., Malyankar,U.M.,
Kekuda,R., Stone,D.J., Anderson,D., Shimkets,R.A., Burgess,C.E.,
Zerhusen,B.D., Liu,X., Spytek,K.A., Casman,S.J., Boldog,F.L.,
Smithson,G., Li,L. and Ji,W.
TITLE Polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 02057453-A 13 25-JUL-2002;
Curagen Corporation (US)
FEATURES
source Location/Qualifiers
1..8354
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 7.51e-33 Length: 8354
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AX556500 (1-8354)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||||
Db 7982 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 8041
|||||

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
|||||

Db 3488 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 3547

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40

Db 3548 TACGGGCACCTGTCTTGAACACACAGCTACGGGACAACCTTGGATGAGGAGGCACGG 3607

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60

Db 3608 GTCCTGGAGCTGGCCCGCAGAGAGCCGTGCGCCAAGCGTGGCCCGCAGCAGACAGA 3667

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76

Db 3668 CTGCGGGAAGGGAGGAGGCTGCGGCCTGGACAGAGGGGAGAAG 3715

RESULT 7

AX600210

LOCUS AX600210

DEFINITION Sequence 22 from Patent WO02072830.

ACCESSION AX600210

VERSION AX600210.1 GI:28400252

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Yue,H., Yao,M.G., Ison,C.H., Lu,Y., Warren,B.A., Elliott,V.S., Baughn,M.R., Ding,L., Xu,Y., Gietzen,K.J., Tang,T.Y., Lal,P.G., Duggan,B.M., Burford,N., Lu,D.A., Richardson,T.W., Tran,U.K., Khare,R. and Walia,N.K.

TITLE Proteins associated with cell growth, differentiation, and death

JOURNAL Patent: WO 02072830-A 22 19-SEP-2002;

INCYTE Genomics, Inc. (US)

FEATURES

source

1. .8645

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

/note="Incyte ID No: 7488573CB1"

ORIGIN

Alignment Scores:

Pred. No.: 7.79e-33 Length: 8645

Score: 391.00 Matches: 76

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AX600210 (1-8645)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20

Db 8031 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 8090

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40

Db 8091 TACGGGCACCTGTCTTGAACACACAGCTACGGGACAACCTTGGATGAGGAGGCACGG 8150

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60

Db 8151 GTCCTGGAGCTGGCCCGCAGACAGCCGTGCGCCAAGCGTGGCCCGCAGCAGACAGA 8210

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76

Db 8211 CTGCGGGAAGGGAGGAGGCTGCGGCCTGGACAGAGGGGAGAAG 8258

RESULT 8

HSM806812

LOCUS HSM806812

DEFINITION Homo sapiens mRNA; cDNA DKFZp686K11107 (from clone DKFZp686K11107).

ACCESSION BX640737

VERSION BX640737.1 GI:34364828

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 8993)

AUTHORS Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.

CONSRITM The German Human cDNA Consortium

TITLE Direct Submission

JOURNAL Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKFZp686K11107) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/.

FEATURES

Location/Qualifiers

1. .8993

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/map="11q14.1"

/clone="DKFZp686K11107"

/tissue_type="human fetal kidney"

/clone_lib="686 (synonym: hlcc3)". Vector pSport1_sfi; host DH10B; sites SfiIA + SfiIB"

/dev_stage="fetal"

1. .8993

/gene="DKFZp686K11107"

<1. .4200

/gene="DKFZp686K11107"

/note="hypothetical protein, N-terminus elongated, differentially spliced"

/codon_start=1

/product="hypothetical protein"

/protein_id="CAE45850.1"

/db_xref="GI:34364829"

/translation="RIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVHLEWPTDLAI NPMDNSLYVLNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESAIA LAVSHNGVLYIAETDEKKINRIQVTTSGEISLVAGAPSGCDCKNDANCDCFSGDDGY AKDAKLTNPSSLAVCADGELYVADLGNIRIRIRKKNKPFLLNTQNMYSLSPIQELYL FDTTGKLYTQSLPTGDLYNFTYTGDDITLITDNNGNMNVRRDSTGMPLWLVPD QVYVWTMGTNSALKSVTTQGHELAMTYHGNSGLLATKSNENGWTTTFEYDSFGRLT NVTFTQGVSSFRSDTSSVHVQVETSSKDDVTITTNLSASGAFYTLQDVRNSYYI GADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGRNVTLPIDNGLNLEVRQRKEQA RGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFTLRILYDQAGRSLWSPSS RLNGVNTYSPGGYIAGIQRGIMSERMEYDQAGRITSRIFADGKTWSYTYLEKSMVLL LHSQROYIFEEDKNDRLSSVTMPNVARQILETIRSVGYRNIYQPPEGNASVIQDFTTE DGHLLHTFYLGTGRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGMLKTINLQNEGFT CTIRYRQIGPLIDRQIFRFTTEGMVNARFDYNDNSFRVTSMQAVINETPLPIDLYRY DDVSGKTEQFGKFGVIYDINQIITTAVMTHTKHFDAYGRMKEVQYEIFRSLMYWMTV QYDNMGRVVKELKVGPYANTTRYSYEYDADGQLQTVSINDKPLWRYSYDLNGNLHL SPGNSARLTPLRDIIRDRIIRLGDVQYKMDGDFLRQGGDI FEYNSAGLLIKAYNRA GSWSVRYRYDGLGRRVSSKSSHHLQFFYADLTNPTKVTHLYNHSSEITSLYDYLQ GHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYATAGEIYMDTDPNFQIII GYHGGLYDPLTKLVHGMRRDYDLAGRWTSPDHKLWKLSSSNVMPFNLYMFKNNNPI SNSQDIKCMTDVSNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDNS KSILGVQCEVQKQLKAFVTLERFDQLYGSTITSQQAPKTKKFASGSVFGKGVKFAI KDRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPSPSEGDLAIL GLSGGRTLENGVNVTVSQINTVNLGRTRRYTDIQLQYGALCLNTRYGTTLDEEKARV LELARQAVRQAWAREQQRLREGEEGLRAWTEGKQQVLSTGRVQGYDGFVVISVEQY PELSDSANNIHFMRQSEMGR"

8953. .8958

/gene="DKFZp686K11107"

8973

polyA_signal

polyA_site

/gene="DKFZp686K11107"

ORIGIN

Alignment Scores: 8.14e-33 Length: 8993
Pred. No.: 391.00 Matches: 76
Score: 391.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 9

US-10-029-020-14_COPY_2650_2725 (1-76) x HSM806812 (1-8993)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 3838 TCCCGATCAACACAGTACTTATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 3897
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 3898 TACGGGGCACTGTGCTTGAACACACACGCTACGGGACACACGTTGGATGAGGAGGACCG 3957
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 3958 GTCCTGGAGTGGCCCGGCAGAGAGCCGTGCGCAAGCGTGGGCCCGGAGCAGCAGAGA 4017
QY 61 LeuArgGluGluGluGluGluLeuArgAlaTrpThrGluGlyGluLys 76
Db 4018 CTGCGGGAAGGGAGGAAGGCGCTGCGGCGCTGGACAGAGGGGGGAGAG 4065

RESULT 9

AC090378/c
LOCUS AC090378 68970 bp DNA linear HTG 17-FEB-2001
DEFINITION Homo sapiens chromosome 18 clone RP11-886F14 map 18, LOW-PASS
SEQUENCE SAMPLING.

AC090378

VERSION AC090378.1 GI:12958014

KEYWORDS HTG; HTGS_PHASE0.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 68970)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 18, clone RP11-886F14.

Unpublished

2 (bases 1 to 68970)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L12672
Center clone name: 886_F_14

* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 714: contig of 714 bp in length
715 814: gap of 100 bp
815 1552: contig of 738 bp in length
1553 1652: gap of 100 bp
1653 2377: contig of 725 bp in length
2378 2477: gap of 100 bp
2478 3210: contig of 733 bp in length
3211 3310: gap of 100 bp
3311 4036: contig of 726 bp in length
4037 4136: gap of 100 bp
4138 4837: contig of 701 bp in length
4838 4937: gap of 100 bp
4938 5678: contig of 741 bp in length
5679 5778: gap of 100 bp
5779 6473: contig of 695 bp in length
6474 6573: gap of 100 bp
6574 7277: contig of 704 bp in length
7278 7377: gap of 100 bp
7378 8120: contig of 743 bp in length
8121 8220: gap of 100 bp
8221 8965: contig of 745 bp in length
8966 9065: gap of 100 bp
9066 9798: contig of 733 bp in length
9799 10591: contig of 693 bp in length
10592 10691: gap of 100 bp
10692 11413: contig of 722 bp in length
11414 11513: gap of 100 bp
11514 12232: contig of 719 bp in length
12233 12332: gap of 100 bp
12333 13067: contig of 735 bp in length
13068 13167: gap of 100 bp
13168 13898: contig of 731 bp in length
13899 13998: gap of 100 bp
13999 14707: contig of 709 bp in length
14708 14807: gap of 100 bp
14808 15514: contig of 707 bp in length
15515 15614: gap of 100 bp
15615 16313: contig of 699 bp in length
16314 16413: gap of 100 bp
16414 17161: contig of 748 bp in length
17162 17261: gap of 100 bp
17262 17981: contig of 720 bp in length
17982 18081: gap of 100 bp
18082 18803: contig of 722 bp in length
18804 18903: gap of 100 bp
18904 19632: contig of 729 bp in length
19633 19732: gap of 100 bp
19733 20441: contig of 709 bp in length
20442 20541: gap of 100 bp
20542 21272: contig of 731 bp in length
21273 21372: gap of 100 bp
21373 22089: contig of 717 bp in length

* 22090 22189: gap of 100 bp
* 22190 22934: contig of 745 bp in length
* 22935 23034: gap of 100 bp
* 23035 23777: contig of 743 bp in length
* 23778 23877: gap of 100 bp
* 23878 24592: contig of 715 bp in length
* 24593 24692: gap of 100 bp
* 24693 25410: contig of 718 bp in length
* 25411 25510: gap of 100 bp
* 25511 26204: contig of 694 bp in length
* 26205 26304: gap of 100 bp
* 26305 27012: contig of 708 bp in length
* 27013 27112: gap of 100 bp
* 27113 27850: contig of 738 bp in length
* 27851 27950: gap of 100 bp
* 27951 28668: contig of 718 bp in length
* 28669 28768: gap of 100 bp
* 28769 29516: contig of 748 bp in length
* 29517 29616: gap of 100 bp
* 29617 30314: contig of 698 bp in length
* 30315 30414: gap of 100 bp
* 30415 31159: contig of 745 bp in length
* 31160 31259: gap of 100 bp
* 31260 31979: contig of 720 bp in length
* 31980 32079: gap of 100 bp
* 32080 32806: contig of 727 bp in length
* 32807 32906: gap of 100 bp
* 32907 33607: contig of 701 bp in length
* 33608 33707: gap of 100 bp
* 33708 34398: contig of 691 bp in length
* 34399 34498: gap of 100 bp
* 34499 35243: contig of 745 bp in length
* 35244 35343: gap of 100 bp
* 35344 36072: contig of 729 bp in length
* 36073 36172: gap of 100 bp
* 36173 36899: contig of 727 bp in length
* 36900 36999: gap of 100 bp
* 37000 37726: contig of 727 bp in length
* 37727 37826: gap of 100 bp
* 37827 38538: contig of 712 bp in length
* 38539 38638: gap of 100 bp
* 38639 39376: contig of 738 bp in length
* 39377 39476: gap of 100 bp
* 39477 40212: contig of 736 bp in length
* 40213 40312: gap of 100 bp
* 40313 41067: contig of 755 bp in length
* 41068 41167: gap of 100 bp
* 41168 41870: contig of 703 bp in length
* 41871 41970: gap of 100 bp
* 41971 42687: contig of 717 bp in length
* 42688 42787: gap of 100 bp
* 42788 43479: contig of 692 bp in length
* 43480 43579: gap of 100 bp
* 43580 44304: contig of 725 bp in length
* 44305 44404: gap of 100 bp
* 44405 45130: contig of 726 bp in length
* 45131 45230: gap of 100 bp
* 45231 45943: contig of 713 bp in length
* 45944 46043: gap of 100 bp
* 46044 46764: contig of 721 bp in length
* 46765 46864: gap of 100 bp
* 46865 47591: contig of 727 bp in length
* 47592 47691: gap of 100 bp
* 47692 48417: contig of 726 bp in length
* 48418 48517: gap of 100 bp
* 48518 49216: contig of 699 bp in length
* 49217 49316: gap of 100 bp
* 49317 50042: contig of 726 bp in length
* 50043 50142: gap of 100 bp
* 50143 50882: contig of 740 bp in length
* 50883 50982: gap of 100 bp
* 50983 51700: contig of 718 bp in length
* 51701 51800: gap of 100 bp

* 51801 52548: contig of 748 bp in length
* 52549 52648: gap of 100 bp
* 52649 53362: contig of 714 bp in length
* 53363 53462: gap of 100 bp
* 53463 54172: contig of 710 bp in length
* 54173 54272: gap of 100 bp
* 54273 54997: contig of 725 bp in length
* 54998 55097: gap of 100 bp
* 55098 55817: contig of 720 bp in length
* 55818 55917: gap of 100 bp
* 55918 56644: contig of 727 bp in length
* 56645 56744: gap of 100 bp

Alignment Scores:

Pred. No.: 7.46e-32 Length: 68970
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AC090378 (1-68970)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||
Db 42327 TCCCAGATCAACACAGTACTTATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 42268

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluYsAlaArg 40
|||
Db 42267 TACGGGGCACTGTGCTTGAACACACACGCTACGGGACCAACGTTGGATGAGGAGGACCGG 42208

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
|||
Db 42207 GTCCTGGAGCTGGCCCGCAGAGAGCGCTGCGCCAAAGCGTGGGCCCGGACGACGAGA 42148

QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrThrGluGlyGluYs 76
|||
Db 42147 CTGCGGGAGGGGAGGAGGCGCTGCGGCCTGGACAGAGGGGAGAGAG 42100

RESULT 10
AP002412/c

LOCUS

DEFINITION Homo sapiens chromosome 11 clone RP11-7H7 map 11q, WORKING DRAFT
SEQUENCE, 18 unordered pieces.

ACCESSION AP002412

VERSION AP002412.1 GI:8131676

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 170156)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 170,156 genomic DNA of 11q

Published Only in DataBase (2000)

REFERENCE 2 (bases 1 to 170156)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,

Japan (E-mail:hattori@gsc.riken.go.jp)

URL:http://hgp.gsc.riken.go.jp/

Fax:81-42-778-9924

COMMENT

----- Genome Center

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: HumDraft11

* 161168 161267: gap of 100 bp
* 161268 164102: contig of 2835 bp in length
* 164103 164202: gap of 100 bp
* 164203 167603: contig of 3401 bp in length
* 167604 167703: gap of 100 bp
* 167704 170156: contig of 2453 bp in length.

FEATURES

source
1. 170156
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-7H7"
1. 29785
/note="assembly_fragment"
29886. 52941
/note="assembly_fragment"
53042. 71728
/note="assembly_fragment"
71829. 87672
/note="assembly_fragment"
87773. 102073
/note="assembly_fragment"
102174. 109422
/note="assembly_fragment"
109523. 117422
/note="assembly_fragment"
117523. 124911
/note="assembly_fragment clone_end:T7 vector_side:left"
125012. 131158
/note="assembly_fragment clone_end:SP6 vector_side:left"
131259. 135747
/note="assembly_fragment"
135848. 141547
/note="assembly_fragment"
141648. 147053
/note="assembly_fragment"
147154. 151529
/note="assembly_fragment"
151630. 156372
/note="assembly_fragment"
156473. 161167
/note="assembly_fragment"
161268. 164102
/note="assembly_fragment"
164203. 167603
/note="assembly_fragment"
167704. 170156
/note="assembly_fragment"

ORIGIN

Alignment Scores: 1.99e-31 Length: 170156
Pred. No.: 391.00 Matches: 76
Score: 391.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2

US-10-029-020-14_COPY_2650_2725 (1-76) x AP002412 (1-170156)

Qy 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 116153 TCCAGATCAACACAGTACTTATGGCAGGACTAGACGCTACACAGACATCCAGTCCAG 116094
Qy 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluLysAlaArg 40
Db 116093 TACGGGCACTGTGCTTGAACACACACGCTACGGGACAACCGTTGGATGAGGAGGACACGG 116034
Qy 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGluGlnArg 60
Db 116033 GTCTGGAGCTGGCCCGCAGAGAGCCGTGGCCCAAGCGTGGGCCCCGAGCAGCAGAGA 115974

Center clone name: RP11-7H7
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 159074 bases at least Q40
Consensus quality: 164580 bases at least Q30
Consensus quality: 167092 bases at least Q20
Insert size: 168456; sum-of-contigs
Quality coverage: 5.57x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 29785 contig of 29785 bp in length
29886 52941 contig of 23056 bp in length
53042 71728 contig of 18687 bp in length
71829 87672 contig of 15844 bp in length
87773 102073 contig of 14301 bp in length
102174 109422 contig of 7249 bp in length
109523 117422 contig of 7900 bp in length
117523 124911 contig of 7389 bp in length
125012 131158 contig of 6147 bp in length
131259 135747 contig of 4489 bp in length
135848 141547 contig of 5700 bp in length
141648 147053 contig of 5406 bp in length
147154 151529 contig of 4376 bp in length
151630 156372 contig of 4743 bp in length
156473 161167 contig of 4695 bp in length
161268 164102 contig of 2835 bp in length
164203 167603 contig of 3401 bp in length
167704 170156 contig of 2453 bp in length.

* NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 29785: contig of 29785 bp in length
* 29786 29885: gap of 100 bp
* 29886 52941: contig of 23056 bp in length
* 52942 53041: gap of 100 bp
* 53042 71728: contig of 18687 bp in length
* 71729 71828: gap of 100 bp
* 71829 87672: contig of 15844 bp in length
* 87673 87772: gap of 100 bp
* 87773 102073: contig of 14301 bp in length
* 102074 102173: gap of 100 bp
* 102174 109422: contig of 7249 bp in length
* 109423 109522: gap of 100 bp
* 109523 117422: contig of 7900 bp in length
* 117423 117522: gap of 100 bp
* 117523 124911: contig of 7389 bp in length
* 124912 125011: gap of 100 bp
* 125012 131158: contig of 6147 bp in length
* 131159 131258: gap of 100 bp
* 131259 135747: contig of 4489 bp in length
* 135748 135847: gap of 100 bp
* 135848 141547: contig of 5700 bp in length
* 141548 141647: gap of 100 bp
* 141648 147053: contig of 5406 bp in length
* 147054 147152: gap of 100 bp
* 147154 151529: contig of 4376 bp in length
* 151530 151629: gap of 100 bp
* 151630 156372: contig of 4743 bp in length
* 156373 156472: gap of 100 bp
* 156473 161167: contig of 4695 bp in length

QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 115973 CTGCGGAAGGGAGGAGGCTGCGGCCTGGACAGAGGGGAGAAG 115926

RESULT 11
AP002768/c
LOCUS AP002768 186084 bp DNA linear PRI 15-MAR-2003
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-673F18,
complete sequence.

ACCESSION AP002768
VERSION AP002768.3 GI:13429926
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens genomic DNA
JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 186084)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Mar 21, 2001 this sequence version replaced gi:11071944.

FEATURES
source
Location/Qualifiers
1..186084
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-673F18"

ORIGIN
Alignment Scores:
Pred. No.: 2.2e-31 Length: 186084
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AP002768 (1-186084)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 27964 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 27905

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluLysAlaArg 40
Db 27904 TAGGGGCACCTGTGCTTGAACACACAGCTACGGGACAAACGTGGATGAGGAGGACGG 27845

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 27844 GTCCTGGAGCTGGCCCGGCAGAGAGCCGTGCGCCCAAGCGTGGCCCGCAGCAGCAGAGA 27785

QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 27784 CTGCGGAAGGGAGGAGGCCTGCGGCCTGGACAGAGGGGAGAAG 27737

RESULT 12
AK056531
LOCUS AK056531 2685 bp mRNA linear PRI 01-AUG-2002
DEFINITION Homo sapiens cDNA FLJ31969 fis, clone NT2RP7008013, highly similar
to Mus musculus mRNA for Ten-m4.

ACCESSION AK056531
VERSION AK056531.1 GI:16551956
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S.,
Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T.,
Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M.,
Kikuchi,H., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2685)
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES
source
Location/Qualifiers
1..2685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP7008013"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP7"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells after 5-weeks retinoic acid (RA)
induction."
242..2434
/note="unnamed protein product"
/protein_id="BAB71206.1"
/db_xref="GI:16551957"
/translation="MLKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTTEGMVNARFDY
NYDNSFRVTSMQAVINETPLPIDLYRYDDVSGTEQFGKFGVIYDINQIITAVMTH
TKHFDAYGRMKEVQYEIFRSLMYWMTVQYDNMGVVKELKVGYPYANTTRYSEYDAD
GQLQTVSINDKPLWRYSYDLNGNLHLLSPGNSARLTPLRYDIRDRIIRLGDVQYKMD
DGFLRQGGDIFEYNSAGLLIKAYNRAGSVSVRYDYDGLGRRVSSKSSHSHLQFFYA
DLTNPTKVTHLYNHSSEITSLYYDLQHLFAMELSSGDEFYIACDNIGTPLAVFSGT
GLMIKQILYTAYGEIYMDTNPNFQIIIGVHGGLYDPLTKLVHMGRRDYDVLGRWTSP
DHELWKHLSSSNVMPFNLYMFKNNPISNSQDIKCFMTDVSNSWLLTTFGQLHNVIPGY
PKPDMAMEPSYELIHTQMKTQEWDNKSILGVQCEVQKQKAFVTLERFDQLYGSTI
TSCLOAPKYKFPASSGSVFGKGVKFKALDGRVTTDIISVANEDGRRVAATLNHAHYLE
NLHFTIDGVTDTHYFVKPGPSEGDAILLGSGRRTLENGNVTVSQINTVLSGRTRRY
TDIQLQYGALCLNTRYGTTLDEEKARVLELARQAVRQAWAREQQQRREGEGLRAWT
EGEKQQVLSTGRVQGYDGFVVISVEQPELSDSANNIHFMRSQSEMGRR"

ORIGIN
Alignment Scores:
Pred. No.: 7.74e-33 Length: 2685
Score: 386.00 Matches: 75
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.68% Mismatches: 0
Query Match: 98.72% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AK056531 (1-2685)			US-10-029-020-14_COPY_2650_2725 (1-76) x GGA293019 (1-834)		
QY	1	SerGlnIleAsnThrValIleuAsnGlyArgThrArgArgTyrThrAspIleGlnIleuGln 20	QY	1	SerGlnIleAsnThrValIleuAsnGlyArgThrArgArgTyrThrAspIleGlnIleuGln 20
Db	2072	TCCAGATCAACACAGTACTAGTGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 2131	Db	550	TCCAGATCAACACAGTCTGGTGGAGGACTAGGCGTTACACGGACATCCAGCTGCAG 609
QY	21	TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrIleuAspGluGluLysAlaArg 40	QY	21	TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrIleuAspGluGluLysAlaArg 40
Db	2132	TACGGGCACTGTGCTTGAACACACAGCTACGGGACACCTGCGGACACCTGGATGAGGAGGACACGG 2191	Db	610	TATGGTGGCTGTGCTGAACACTCGTACGGGACCACTTTGGACGAGGAGAGGCCGGA 669
QY	41	ValIleuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60	QY	41	ValIleuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db	2192	GTCCTGGAGCTGGCCCCGGCAGAGAGCCGTGCGCCAAGCTGGGCCCGCGCAGCAGCAGAGA 2251	Db	670	GTCTCTGGAGCTGGCCCCGGCAGAGCGCGCTGGCCCAAGCTTGGGCCCGGGAACAGCAGAGA 729
QY	61	LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76	QY	61	LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db	2252	CTGCGGAAGGGAGGAAGGCCTGCGGGCCTGGACAGAGGGGGGAGAAG 2299	Db	730	TTGAGGGATGGGGAGGAGGGGATTCGCTCGTGGACAGAGGGGAGAAG 777
RESULT 13			RESULT 14		
GGA293019			AC074048		
Gallus gallus partial mRNA for teneurin-4 (TEN4 gene).			Mus musculus chromosome 7 clone RP23-18J2, WORKING DRAFT SEQUENCE,		
AJ293019			6 unordered pieces.		
AJ293019.1 GI:9909616			AC074048		
TEN4 gene; teneurin-4.			AC074048.17 GI:14547771		
Gallus gallus (chicken)			HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
Gallus gallus			Mus musculus (house mouse)		
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;			Mus musculus		
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Phasianinae; Gallus.			Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
1 Tucker, R.P., Martin, D., Kos, R. and Chiquet-Ehrismann, R.			1 (bases 1 to 208760)		
The expression of teneurin-4 in the avian embryo			Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,		
Mech. Dev. 98 (1-2), 187-191 (2000)			Dederich, D., Thomas, S., Okwuonu, G., Carlock, C., Garner, T.,		
20500912			Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,		
11044628			Buhay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,		
2 (bases 1 to 834)			Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,		
Chiquet-Ehrismann, R.			Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R.,		
Direct Submission			Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogues, M.,		
Submitted (22-AUG-2000) Chiquet-Ehrismann R., Cell Biology,			Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,		
Friedrich Miescher Institute, Maulbeerstrasse 45, CH-4058 Basel,			Kovar, C., Liu, J., Liu, W., Loulseged, H., Lozado, R.J., Martin, R.,		
SWITZERLAND			Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,		
Location/Qualifiers			Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Oguh, M., Parish, B.,		
1. .834			Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S.,		
/organism="Gallus gallus"			Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A.,		
/mol_type="mRNA"			Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G.,		
/db_xref="taxon:9031"			Worley, K. and Gibbs, R.		
/tissue_lib="11 day embryonic brain"			Direct Submission		
1. .834			Unpublished		
/gene="TEN4"			2 (bases 1 to 208760)		
<1. .>834			Worley, K.C.		
/gene="TEN4"			Direct Submission		
/codon_start=1			Submitted (13-JUL-2000) Human Genome Sequencing Center, Department		
/evidence=experimental			of Molecular and Human Genetics, Baylor College of Medicine, One		
/product="teneurin-4"			Baylor Plaza, Houston, TX 77030, USA		
/protein_id="CAC04511.1"			On Jun 25, 2001 this sequence version replaced gi:13162473.		
/db_xref="GI:9909617"			----- Genome Center		
/db_xref="SPTREMBL:Q9DEQ8"			Center: Baylor College of Medicine		
/translation="YMTDVNSWLLTFGFQLHNVIPGPKPDLDAPEPSYELIHTQMKT			Center code: BCM		
QEWDSKSIILGVQCEVQQLKAFVTLERFEQIYSSSIAGCQQVKKNFASGGSIFGX			Web site: http://www.hgsc.bcm.tmc.edu/		
GVKFAMKDRVATDIISVANEDGRRIIAAILNNAHYLENLHFTIDGVDTHYFIKQGPSE			Contact: hgsc-help@bcm.tmc.edu		
GDLISILGSLGGRRTLEGNVTVSQINTVLGGRTRYTDIQLQYGALCLNTRYGTLLD			----- Project Information		
BEKARVLELARQRAVAQAWAREQQLRDGEEGIRSWTEGEKQQLVLTGRVQGYDGYFV			Center project name: MAIM		
IS"			Center clone name: RP23-18J2		
ORIGIN			----- Summary Statistics		
Alignment Scores:			Sequencing vector: M13; L08821		
Pred. No.:	9.66e-32	Length: 834	Chemistry: Dye-primer Bodipy: 63% of reads		
Score:	371.00	Matches: 71	Chemistry: Dye-terminator Big Dye: 37% of reads		
Percent Similarity:	97.37%	Conservative: 3	Assembly program: Phrap; version 0.990329		
Best Local Similarity:	93.42%	Mismatches: 2	Consensus quality: 210214 bases at least Q40		
Query Match:	94.88%	Indels: 0	Consensus quality: 213555 bases at least Q30		
DB:	5	Gaps: 0	Consensus quality: 215980 bases at least Q20		
			Estimated insert size: 208744; sum-of-contigs estimation		

Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 8.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

- * 1 103064: contig of 103064 bp in length
- * 103065 103164: gap of unknown length
- * 103165 124754: contig of 21590 bp in length
- * 124755 124854: gap of unknown length
- * 124855 147843: contig of 22989 bp in length
- * 147844 147943: gap of unknown length
- * 147944 168730: contig of 20787 bp in length
- * 168731 168830: gap of unknown length
- * 168831 192230: contig of 23400 bp in length
- * 192231 192330: gap of unknown length
- * 192331 208760: contig of 16430 bp in length.

FEATURES

source
1..208760
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/clone="RP23-18J2"

ORIGIN

Alignment Scores:

Pred. No.: 8.38e-29 Length: 208760
Score: 368.00 Matches: 71
Percent Similarity: 97.37% Conservative: 3
Best Local Similarity: 93.42% Mismatches: 2
Query Match: 94.12% Indels: 0
DB: 2 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AC074048 (1-208760)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 4946 TCCAGATCAACACCGTGCTCAGTGGCAGGACTAGACGTACACTGCATCCAGCTGCAG 5005
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 5006 TACAGGGCGCTGTGCCTGAACACCCGCTACGGGACACACAGTGGATGAGGAAAGTGC GG 5065
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 5066 GTGCTGGAGCTGGCCAGGCAGAGAGCTGTGCGCCAGCGCTGGGCCCGAGAGCAGAGAGA 5125
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 5126 CTGCGGGAAGGGAGAGGGCCTGCGGGCCTGGACAGATGGGGAAG 5173

RESULT 15

AC073599/c

LOCUS

DEFINITION Mus musculus 7 BAC RP23-8M3 (Roswell Park Cancer Institute Mouse
BAC Library) complete sequence.

ACCESSION AC073599

VERSION AC073599.16 GI:13162472

KEYWORDS HTG.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 226256)

REFERENCE

AUTHORS Metzker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C.,

Dederich,D., Thomas,S., Okwuonu,G., Carlock,C., Garner,T.,
Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J.,
BuhaY,C., Bunac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z.,
Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S.,
Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R.,
Gorrell,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogues,M.,
Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S.,
Kovar,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Martin,R.,
Massey,E., McLeod,M.P., Mei,G., Moore,S., Morgan,M., Morris,S.,
Neal,D., Nelson,A., Nguyen,R., Nguyen,N., Oguh,M., Parish,B.,
Perez,L., Reiter,D., Say,J., Shen,H., Vasquez,L., Watlington,S.,
Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A.,
Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Weinstock,G.,
Worley,K. and Gibbs,R.

TITLE Direct Submission

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (26-JUN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Feb 28, 2001 this sequence version replaced gi:12963002.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by
Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui
Zhang.

Exon/Intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.

FEATURES

source

1..226256
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="7"
/clone="RP23-8M3"
complement(147..670)
/note="Region similar to Mm#S1670873 601761001F1 Mus
musculus cDNA: BF124644"

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

complement(5737..5848)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 19:26:18 ; Search time 602 Seconds
(without alignments)
3260.204 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76
Perfect score: 2132
Sequence: 1 MDVKERKPYRSLTRRRDAER.....EITEDTASSWPVPTDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/MITRA020/runat_06082004_114103_258/app_query.fasta_1.583
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=MITRA020@cgn_1_1_519@runat_06082004_114103_258
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

1	2127	99.8	8354	13	US-10-383-201-43	Sequence 43, Appl
2	2127	99.8	8354	13	US-10-029-020-13	Sequence 13, Appl
3	2122	99.5	1680	15	US-10-290-578-3	Sequence 3, Appli
4	2122	99.5	3111	9	US-09-773-517-12	Sequence 12, Appl
5	2122	99.5	3111	9	US-09-792-025-12	Sequence 12, Appl
6	2122	99.5	3111	9	US-09-849-868-12	Sequence 12, Appl
7	2122	99.5	3111	15	US-10-290-578-1	Sequence 1, Appli
8	2122	99.5	3111	15	US-10-453-183-12	Sequence 12, Appl
9	2114	99.2	8355	13	US-10-383-201-55	Sequence 55, Appl
10	2082	97.7	8438	13	US-10-042-865-1	Sequence 1, Appli
11	1285	60.3	2387	15	US-10-290-578-11	Sequence 11, Appl
12	975	45.7	8645	17	US-10-038-854-37	Sequence 37, Appl
13	975	45.7	8675	17	US-10-038-854-35	Sequence 35, Appl
14	937	43.9	9695	16	US-10-144-194A-81	Sequence 81, Appl
15	905	42.4	2496	9	US-09-808-602-75	Sequence 75, Appl
16	905	42.4	2496	10	US-09-800-198-63	Sequence 63, Appl
17	904	42.4	8689	9	US-09-808-602-78	Sequence 78, Appl
18	904	42.4	8689	10	US-09-800-198-66	Sequence 66, Appl
19	904	42.4	8797	9	US-09-808-602-74	Sequence 74, Appl
20	904	42.4	8797	9	US-09-808-602-77	Sequence 77, Appl
21	904	42.4	8797	10	US-09-800-198-62	Sequence 62, Appl
22	904	42.4	8797	10	US-09-800-198-65	Sequence 65, Appl
23	885	41.5	8409	9	US-09-808-602-79	Sequence 79, Appl
24	885	41.5	8409	10	US-09-800-198-67	Sequence 67, Appl
25	859	40.3	8575	13	US-10-072-012-143	Sequence 143, App
26	828	38.8	9729	9	US-09-808-602-12	Sequence 12, Appl
27	828	38.8	9729	10	US-09-800-198-12	Sequence 12, Appl
28	828	38.8	9826	9	US-09-808-602-7	Sequence 7, Appli
29	828	38.8	9826	10	US-09-800-198-7	Sequence 7, Appli
30	733.5	34.4	12880	16	US-10-295-027-927	Sequence 927, App
31	641.5	30.1	8473	17	US-10-038-854-39	Sequence 39, Appl
32	641.5	30.1	8487	17	US-10-038-854-41	Sequence 41, Appl
33	522	24.5	1755	10	US-09-998-966-17	Sequence 17, Appl
34	522	24.5	1755	15	US-10-004-415-17	Sequence 17, Appl
35	522	24.5	1755	16	US-10-384-974-17	Sequence 17, Appl
36	508	23.8	1727	16	US-10-094-749-710	Sequence 710, App
37	508	23.8	9058	16	US-10-144-194A-79	Sequence 79, Appl
38	452	21.2	527	15	US-10-029-386-1356	Sequence 1356, Ap
39	434	20.4	1430	9	US-09-808-602-5	Sequence 5, Appli
40	434	20.4	1430	10	US-09-800-198-5	Sequence 5, Appli
41	434	20.4	1431	10	US-09-977-418-21	Sequence 21, Appl
42	434	20.4	1431	10	US-09-977-033A-21	Sequence 21, Appl
43	434	20.4	1431	10	US-09-977-751C-21	Sequence 21, Appl
44	434	20.4	1431	10	US-09-977-639A-21	Sequence 21, Appl
45	434	20.4	1431	11	US-09-977-819B-21	Sequence 21, Appl

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:

Pred. No.: 5.86e-206 Length: 8354
Score: 2127.00 Matches: 397
Percent Similarity: 99.25% Conservative: 0
Best Local Similarity: 99.25% Mismatches: 3
Query Match: 99.77% Indels: 0
DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-10-383-201-43 (1-8354)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAlaGluArg 20
Db 35 ATGACGTGAAGAGAGGAAGCCCTTACCGCTCGCTGACCCGGCGCGGACCGCGGCGC 94
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 95 CGCTACACCAAGCTCGTCCGCGGACAGCGAGGAGGGCAAGCCCCGGAGAAATCGTACAGC 154
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 155 TCCAGCGAGACCCCTGAAGCCCTACGACGAGACGCCCGCTAGCCTATGGCAGCCGCGTC 214
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPhThrLeu 80
Db 215 AAGGACATTGTGCCGCGAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCAACCTG 274
QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 275 CGGGAGCTGGGGCTGGAAGAAGTAAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 334
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 335 GGCTGCCCCAATGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 394
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 395 ACGTGCTGTCCCTGAGCACCCCGTGTGTGGGGCCGAGCACACGGTCAGGGCGC 454
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 455 AGCTCCTGCTGTCCAGCCGGGCCAAATCCAAATCTCACACTCACCGACACCGAGCATGAA 514
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 515 AACACTGAGACTGATCATCCGGCGGCGCTGCAGAACCAACCGCGGCTCCGGACGCCGCG 574
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 575 CCGCGCTCTCGCACGCCCAACACCCCAACACGACACCGCGGCTCCATTAACCTCCCTG 634
QY 201 AsnArgGlyAsnPhThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 635 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCGCGGCCCCCACGGACCACTCGCTC 694
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240

Db 695 TCCGGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCAGGAACTGGCTG 754
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 755 CTCAACAGCAACATCCCTCCCTGGAGACCAGGAACCTTAGGCAAGCAGCCATTCTTAGGACA 814
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 815 TTGCAGGACAACCTCATTTGAGATGGACATTTCTCGGGCGCTCCCGCATGATGGGCTTAC 874
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 875 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCATCA 934
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 935 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGC 994
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 995 AGCACCTTCGCCCGCGCGGCCTTTAAACCTCAAGAAAGCCCTCCAACTACTGTAACTGGAAG 1054
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1055 TGCGCAGCCCTGAGCGCCATCGTCACTCTCAGCCACTCTGGTCACTCTCTGGCATACTTT 1114
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1115 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1174
QY 381 GluIleThrGluAspThrAlaSerSerThrProValProThrAspValSerLeuTyrPro 400
Db 1175 GAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACCGTCTCCCTATACCCC 1234

RESULT 2

US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:

Pred. No.: 5.86e-206 Length: 8354

Score:		2127.00	Matches:	397
Percent Similarity:		99.25%	Conservative:	0
Best Local Similarity:		99.25%	Mismatches:	3
Query Match:		99.77%	Indels:	0
DB:		13	Gaps:	0
SEQ14-X-AT-28-64-76 (1-400) x US-10-029-020-13 (1-8354)				
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg	20	
Db	35	ATGGACGTGAAGGAGAGGAAGCCCTTACCGCTCGCTGACCCGGCGCGGACGCCGAGCGC	94	
QY	21	ArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyrSer	40	
Db	95	CGCTACACAGCTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCGCAGAAAATCGTACAGC	154	
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60	
Db	155	TCCAGCGAGACCCCTGAAGGCCTACGACGAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC	214	
QY	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80	
Db	215	AAGGACATTGTGCGCGAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG	274	
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100	
Db	275	CGGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT	334	
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120	
Db	335	GGCCTGCCCAATGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGCTGAC	394	
QY	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140	
Db	395	ACGGTGTGTCCCTGAGCACCCCGTGCCTGTGTGGGCGGAGCACACGGTCAGGGCGC	454	
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160	
Db	455	AGCTCTGCCTGTCCAGCGGGCCAAATTCCAATCTCACACTCACGACACCGAGCATGAA	514	
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180	
Db	515	AACACTGAGACTGATCATCCGGGCGGCGCTGCAGAACCAACGCGGGCTCCGGACGCCGCG	574	
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200	
Db	575	CCGCGCTCTCGCACGCCACACACCCCAACACGACCAACCGCGGCTCCATTAACCTCCCTG	634	
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220	
Db	635	AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCACGGACCACTCGCTC	694	
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240	
Db	695	TCCGAGAGCCCCCTGCGGCGGCGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTG	754	
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260	
Db	755	CTCAACAGCAACATCCCCCTGGAGACCAGGAACCTAGGCAAGCAGCCATTCTCTAGGGACA	814	
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280	
Db	815	TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGGCTTAC	874	
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300	
Db	875	AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCGACATCA	934	
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320	
Db	935	CCAGGGTACCCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCTGCCCGC	994	
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340	

Db	995	AGCACCTTCGCCCGCGCCCTTTAACTCAAGAAGCCCTCCAAGTACTGTAAGTGAAG	1054	
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360	
Db	1055	TGCGCAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCATCTCTGGCATACTTT	1114	
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380	
Db	1115	GTGGCCATGCACCTGTTTGGCCTAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT	1174	
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400	
Db	1175	GAGATCAGGAGGACACAGCCAGCAGTGGCCTGTGCCAACCCGACGTCTCCCTATACCCC	1234	
RESULT 3				
US-10-290-578-3				
; Sequence 3, Application US/10290578				
; Publication No. US20030078389A1				
; GENERAL INFORMATION:				
; APPLICANT: Schaefer, Gabriele M.				
; Sliwkowski, Mark				
; TITLE OF INVENTION: Gamma-Herregulin				
; NUMBER OF SEQUENCES: 11				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Genentech, Inc.				
; STREET: 460 Point San Bruno Blvd				
; CITY: South San Francisco				
; STATE: California				
; COUNTRY: USA				
; ZIP: 94080				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: WinPatIn (Genentech)				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/10/290,578				
; FILING DATE: 08-No. US20030078389A1-2002				
; CLASSIFICATION: <Unknown>				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: US/08/891,845				
; FILING DATE: <Unknown>				
; APPLICATION NUMBER: 60/021640				
; FILING DATE: 12-Jul-96				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Lee, Wendy M.				
; REGISTRATION NUMBER: 40,378				
; REFERENCE/DOCKET NUMBER: P1043				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 415/225-1994				
; TELEFAX: 415/952-9881				
; TELEX: 910/371-7168				
; INFORMATION FOR SEQ ID NO: 3:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 1680 base pairs				
; TYPE: Nucleic Acid				
; STRANDEDNESS: Single				
; TOPOLOGY: Linear				
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:				
US-10-290-578-3				
Alignment Scores:				
Pred. No.:		2.23e-206	Length:	1680
Score:		2122.00	Matches:	396
Percent Similarity:		99.00%	Conservative:	0
Best Local Similarity:		99.00%	Mismatches:	4
Query Match:		99.53%	Indels:	0
DB:		15	Gaps:	0
SEQ14-X-AT-28-64-76 (1-400) x US-10-290-578-3 (1-1680)				
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20	

Db 1 ATGGACGTGAAGGAGAGAGCCTTACCGCTCGCTGACCCGGCGCGCAGCGCGAGCGC 60
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 61 CGCTACACCAGCTCGTCCGCGACAGCGAGGAGGCGAAAGCCCGCAGAAATCGTACAGC 120
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 121 TCCAGCGAGACCCCTGAAGCCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGCGTC 180
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 181 AAGGACATTGTCCCGCAGGAGGCCGAGGAATTCTGCCGACAGGTGCCAACTTCACCCCTG 240
QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 241 CGGGAGCTGGGGCTGGAAGATTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 300
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 301 GGCTTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 360
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 361 ACGGTGCTGTCCCTTGAGCACCCCGTGCGTCTGTGGGGCCGGAGCACACGGTTCAGGGCGC 420
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 421 AGCTCTCTGCTGCCAGCGGGCCAAATCCAAATCTCACACTCACGACACCGAGCATGAA 480
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 481 AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCAACCGCGGCTCCGGACGCCGCCG 540
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 541 CCGCCGCTCTCGCACGCCCACACCCCAACACGACCAACCGCGGCTCCATTAACTCCCTG 600
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 601 AACCGGGGCAACTTCACGCCGAGGAGCAACCCACGCCCGCCCGCCCAACCGACCACTCGCTC 660
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 661 TCCGGAGAGCCCTGCGGGCGGCGCCAGGAGCCTGCCACGCCCGCCAGGAACTGGCTG 720
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 721 CTCAACAGCAACATCCCTTGAGATGGACATTCTGGCGCCTCCCGCCATGATGGGCTTAC 780
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 781 TTGCAGGACAACCTCATTTGATGGACATTCTGGCGCCTCCCGCCATGATGGGCTTAC 840
QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 841 AGTGACGGGCACTTCCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCATCA 900
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 901 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCCGCCCGCCCGCCCTGCCCGCGC 960
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 961 AGCACCTTCGCCCGCGCGCCTTTAACTCAAGAGCCCTCCAAGTACTGTAACTGGAAG 1020
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1021 TGCGCAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCACTCCCTGCTGGCATACTTT 1080
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

Db 1081 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1140
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1141 GAGATCACGGAGGACACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCC 1200
RESULT 4
US-09-773-517-12
; Sequence 12, Application US/09773517
; Patent No. US20010023241A1
; GENERAL INFORMATION:
; APPLICANT: Sliwkowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/09/773,517
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/243,198
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-773-517-12
Alignment Scores:
Pred. No.: 5.07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 9 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x US-09-773-517-12 (1-3111)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGGAGAGAGAGCCCTTACCGCTCGCTGACCCCGCGCGCAGCGCGC 393
QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCAGCTCGTCCGCGGACAGCGAGGAGGCAAGAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGGCCCTACGACCGAGACGCCCGCTAGCCTATGGCAGCCGCTC 513
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTGTGCCGCGAGGAGGCCGAGGAATTCTGCCGACAGGTGCCAACTTCACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGGAGCTGGGGCTGGAAGAAGTAAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGTATGCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCCGTGGCTGTGTGGGGCCGGAGCACACGGTTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTGCTGTCCAGCCGGGCCAAATTCAAATCTCACACTCACCCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGGCGGCTTCAGAAACCAACCGCGGCTCCGGACGCCCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200

```
Db      874 CCGCGCTCTCGCAGCGCCACACCCCAACAGCACACGCGCTCCATTAACTCCCTG 933
QY      201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db      934 AACCGGGCAACTTCAGCGCGAGGAGCAACCCAGCGCGGCCCCACGACCACTCGCTC 993
QY      221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db      994 TCCGAGAGCCCTGCGCGCGCGCCAGAGCTGCCACGCGCCAGGAGAACTGGCTG 1053
QY      241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db      1054 CTCAACAGCAACATCCCTCGAGACCAGAACTAGGCAAGCAGCCATTCCTAGGGACA 1113
QY      261 LeuGlnAspAsnIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db      1114 TTGCAGGACAACTCATTTGAGATGGACATTTCTGGGGCTTCCGGCCATGATGGGCTTAC 1173
QY      281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db      1174 AGTGACGGGCACTTCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCATCA 1233
QY      301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db      1234 CCAGGGTACCCACTGACGTCCAGCACAGTGACTCTCTCCGCCCGACCCCTGCCCGC 1293
QY      321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db      1294 AGCACCTTCGCGCGCGCGCTTTAACTCAAGAAGCCCTCCAACTACTGTAACTGGAAG 1353
QY      341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db      1354 TGCGCAGCCCTGAGCGCATTGTCATCTCAGCCACTCTGGTCTATCTGCTGGCATATT 1413
QY      361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db      1414 GTGGCCATGCACCTGTTTGGCCATAAAGTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473
QY      381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db      1474 GAGATCAGGAGGACACAGCCAGAGTTGGCCTGTGCCAACCGACGTCTCCCTATACCCC 1533
```

RESULT 5

```
US-09-792-025-12
; Sequence 12, Application US/09792025
; Patent No. US20020042087A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
;               Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/792,025
; FILING DATE: 23-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/020,598
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
```

```
;
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: P1145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: nucleic acid
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-792-025-12
```

```
Alignment Scores:
Pred. No.: 5.07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 9 Gaps: 0
```

SEQ14-X-AT-28-64-76 (1-400) x US-09-792-025-12 (1-3111)

```
QY      1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db      334 ATGGACGTGAAGAGAGAGAAAGCCTTACCGCTCGCTGACCCGCGCGCGAGCGCGC 393
QY      21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db      394 CGTACACACAGCTCGTCCGCGGACAGCGAGGAGGGCAAGCCCGCGAGAAATCGTACAGC 453
QY      41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db      454 TCCAGCGAGACCTGAAGGCCTACGACGAGCGCGCGCTAGCCTATGGCAGCGCGCTC 513
QY      61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db      514 AAGGACATTGTCCCGCAGGAGCGCGAGGAATTTCTCCCGCACAGGTGCCAACTTCACCCCTG 573
QY      81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db      574 CGGAGCTGGGGCTGGAGAAAGTAACGCCCTTACCGGACCTGTACCGGACAGACATT 633
QY      101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db      634 GGCCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGTGGGCGCGGAGCACACGTCAGGCGC 693
QY      121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db      694 ACGTGCTGTCCCTGAGCACCCCGTGTCTGTGGGCGCGGAGCACACGTCAGGCGCGC 753
QY      141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db      754 AGCTCCTGCCTGTCCAGCGGGCCAAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY      161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db      814 AACACTGAGACTGATCATCCGGGCGGCTGCAGAACACCGCGGCTCCGGAGCGCGCGC 873
QY      181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db      874 CCGCGCTCTCGCACGCCCCACACCCCCAACCCAGCACCAACGCGGCTCCATTAACTCCCTG 933
QY      201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db      934 AACCGGGCAACTTCAGCGCGAGGAGCAACCCCGCGCGGCCCCACGACCACTCGCTC 993
QY      221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db      994 TCCGAGAGCCCTGCGCGCGCGCGCGAGGCTGCCACCGCGCGAGAGAACTGGCTG 1053
```



```
QY      241  LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr 260
      |||
Db      1054 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTCTAGGGACA 1113

QY      261  LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
      |||
Db      1114 TTGCAGGACAACTCATTTGAGATGGACATTCTCGGGCGCTCCCGCCATGATGGGGCTTAC 1173

QY      281  SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
      |||
Db      1174 AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCA 1233

QY      301  ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
      |||
Db      1234 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCGCCGACCCCTGCCCCCGC 1293

QY      321  SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
      |||
Db      1294 AGCACCTTCGCGCGCGCGCTTTAAACCTCAAGAAGCCCTCCAAAGTACTGTAACTGGAG 1353

QY      341  CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
      |||
Db      1354 TGGCGAGCCCTGAGCGGCATCGTATCTCAGCCACTCTGGTCACTCTGCTGGCATACTTT 1413

QY      361  ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
      |||
Db      1414 GTGGCCATGCACCTGTTGGCCCTAACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1473

QY      381  GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
      |||
Db      1474 GAGATCAGCGAGGACACAGCCAGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 6
US-09-849-868-12
; Sequence 12, Application US/09849868
; Patent No. US20020081299A1
; GENERAL INFORMATION:
;   APPLICANT: Genentech, Inc.
;   APPLICANT: Gao, Wei-Qiang
;   TITLE OF INVENTION: HAIR CELL DISORDERS
;   FILE REFERENCE: GENENT.035VPC
;   CURRENT APPLICATION NUMBER: US/09/849,868
;   CURRENT FILING DATE: 2001-05-04
;   NUMBER OF SEQ ID NOS: 14
;   SOFTWARE: FastSeq for Windows Version 4.0
;   SEQ ID NO 12
;   LENGTH: 3111
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: (334)...(2637)
US-09-849-868-12

Alignment Scores:
Pred. No.:      5.07e-206      Length:      3111
Score:          2122.00      Matches:      396
Percent Similarity: 99.00%      Conservative: 0
Best Local Similarity: 99.00%      Mismatches: 4
Query Match:      99.53%      Indels:      0
DB:              9          Gaps:      0

SEQ14-X-AT-28-64-76 (1-400) x US-09-849-868-12 (1-3111)

QY      1  MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
      |||
Db      334  ATGGACGTGAAGGAGAGAAAGCCTTACCGCTCGCTGACCCGCGCGCGCGCGCGCGCGCGC 393

QY      21  ArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
      |||
Db      394  CGCTACACACGACTCGTCCGCGGACAGCGAGGAGGCGCAAGCCCCCGCAGAAATCGTACAGC 453

QY      41  SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
```

```
Db      454  TCCAGCGAGACCCCTGAAGCCTACGACCAGGACGCGCGCCTAGCCTATGGCAGCGCGGTC 513

QY      61  LysAspIle**ProGlnGluAlaGluGluPheCysArgThrGly**AsnPheThrLeu 80
      |||
Db      514  AAGGACATTGTGCCCGCAGGAGCGCGAGGAATTCTCGGCACACAGGTGCCAACTTCACCCCTG 573

QY      81  ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
      |||
Db      574  CGGAGCTGGGGCTGGAAAGAAAGTAACGCCCTTCACGGGACCTGTACCGGACAGACATT 633

QY      101  GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
      |||
Db      634  GGCCTCCCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693

QY      121  ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
      |||
Db      694  ACGGTGCTGTCCCTGAGCACCCCGTGTCTGTGGGCGGAGCACACGGTCAGGGCGC 753

QY      141  SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
      |||
Db      754  AGTCTCTGCTGTCCAGCGCGGCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813

QY      161  AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
      |||
Db      814  AACACTGAGACTGATCATCCGGCGGCTGTGCAGAACCCACCGCGGCTCCGACCGCGCGC 873

QY      181  ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
      |||
Db      874  CCGCCGCTCTCGCACGCCACACCCCCAACCCAGCACCCACCGGGCTCCATTAACTCCCTG 933

QY      201  AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
      |||
Db      934  AACCGGGCAACTTTCACGCCGAGGAGCAACCCAGCCCGCCCCCAGCGGACCACTCGCTC 993

QY      221  SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
      |||
Db      994  TCCGGAGAGCCCCCTGCCGGCGGCGCGCAGGAGCCTGCCACCGCCAGGAGAACTGGCTG 1053

QY      241  LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
      |||
Db      1054 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTTAGGCAAGCAGCCATTCTTAGGGACA 1113

QY      261  LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
      |||
Db      1114 TTGCAGGACAACTCATTTGAGATGGACATTCTCGGCGCTCCCGCCATGATGGGCTTAC 1173

QY      281  SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
      |||
Db      1174 AGTGACGGGCACCTTCTCTTCAAGCCTTGAGGACACCTCCCGCTCTTCTGCACCAATCA 1233

QY      301  ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
      |||
Db      1234 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCGCCGACCCCTGCCCCCGC 1293

QY      321  SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
      |||
Db      1294 AGCACCTTCGCGCGCGCGCTTTAAACCTCAAGAAGCCCTCCAAAGTACTGTAACTGGAG 1353

QY      341  CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
      |||
Db      1354 TGGCGAGCCCTGAGCGGCATCGTATCTCAGCCACTCTGGTCACTCTGCTGGCATACTTT 1413

QY      361  ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
      |||
Db      1414 GTGGCCATGCACCTGTTGGCCCTAACTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1473

QY      381  GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
      |||
Db      1474 GAGATCAGCGAGGACACAGCCAGTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 1533

RESULT 7
US-10-290-578-1
```

Sequence 1, Application US/10290578
Publication No. US20030078389A1
GENERAL INFORMATION:
APPLICANT: Schaefer, Gabriele M.
Sliwowski, Mark
TITLE OF INVENTION: Gamma-Herregulin
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/290,578
FILING DATE: 08-No. US20030078389A1-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/891,845
FILING DATE: <Unknown>
APPLICATION NUMBER: 60/021640
FILING DATE: 12-Jul-96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3111 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-290-578-1

Alignment Scores:
Pred. No.: 5.07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 15 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-10-290-578-1 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 334 ATGGACGTGAAGGAGGAGGAGCCTTACCGCTCGCTGACCCGCGCGCGACGCGAGCGC 393
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 394 CGCTACACCAGCTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCGCAGAAATCGTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 454 TCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC 513
QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 514 AAGGACATTGTGCGCGAGGAGCGCGAGGAATTCTGCCGACAGGTGCCAACTTACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100

Db 574 CGGAGCTGGGGTGAAGAAGTAACGCCCCCTCAGGGACCCCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCACTCGGCTACTCCATGGGGGCTGGCTCTGATCCGACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTTGAGCACCCCGTGGCTGTGTGGGCGGAGACACGGTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCCTGCTGTCCAGCGGGCCAAATCCAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGGCGGCTGCAGAACCCAGCACCGCGGCTCCGACGCCGCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCCGCTCTCGCACGCCACACCCCAACCCAGCACCGCGGCTCCATTAATCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCACGCCGAGGAGCAACCCCGCGGCCCGCCAGGACCCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGAGAGCCCCCTGCCGGCGGCGCCAGGAGCCTGCCACGCCCGAGGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCTTCTTAGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGAGGACAACTTCATTGATGGACATTCTCGGCGCTCCCGCCATGATGGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCTCTTCAAGCCTTGAGGCACCTCCCGCTCTTCTGCACCACATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGCCCGCCCGCGC 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACCTTCGCCCGCGCGCTTTAACTCAAGAAAGCCCTCCAAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1354 TGGCAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCTATCCTGTGGCATACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCCATGCACCTGTGTGGCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCACGGAGGACACAGCCAGCAGTGGCCCTGTGCCAACCGACGCTCTCCCTATACCC 1533

RESULT 8
US-10-453-183-12
; Sequence 12, Application US/10453183
; Publication No. US20030199429A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/10/453,183
; CURRENT FILING DATE: 2003-06-03

; PRIOR APPLICATION NUMBER: US/09/243,198
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,866
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-453-183-12

Alignment Scores:
Pred. No.: 5,07e-206 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 15 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-10-453-183-12 (1-3111)

QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg	20
Db	334	ATGGACGTGAAGGAGAGAGACCTTACCGCTCGCTGACCCCGCGCCGACGCCGAGCGC	393
QY	21	ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	394	CGCTACACCAGCTCGTCCGCGGACAGCGGAGGAGGGCAAAGCCCGCGCAGAAATCGTACAG	453
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	454	TCCAGCGAGACCCCTGAAGCCCTACGACCAGGACGCCCGCTAGCTATGGCAGCCCGCTC	513
QY	61	LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu	80
Db	514	AAGGACATTGTGCCGAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG	573
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	574	CGGAGCTGGGGCTGGAAGAAAGTAACGCCCCCTCACGGGACCCTGTACCGGACAGACATT	633
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	634	GGCTCCCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	693
QY	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140
Db	694	ACGGTGTCTCCCTTGAGCACCCCGTGTGTGGGCGCGGAGCACACGGTCAGGGCGC	753
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160
Db	754	AGCTCCTGCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA	813
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180
Db	814	AACACTGAGACTGATCATCCGGCGGCTGCAGAACCCAGCCGGCTCCGACGCCGCCG	873
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu	200
Db	874	CCGCCGCTCTCGACGCCCCACACCCCAACACGACCAACCGGGCTCCATTAACTCCCTG	933
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220
Db	934	AACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCCAACGACCACTCGCTC	993
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240
Db	994	TCCGAGAGCCCCCTGCCGGCGGCCCGGAGGCTGCCACGCCCGGAGGAACTGGCTG	1053
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260
Db	1054	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTCCTAGGGACA	1113

QY	261	LeuGlnAspAsnLeulleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280
Db	1114	TTGCAGGACAACCTCATTTGAGATGGACATTTCTGGCGCCTCCCGCCATGATGGGCTTAC	1173
QY	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300
Db	1174	AGTGACGGGCACCTTCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCATCA	1233
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg	320
Db	1234	CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCCCGACCCCTGCCCGC	1293
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340
Db	1294	AGCACCTTCGCCCGCGCGCTTTAACCTCAAGAAAGCCCTCCAAGTACTGTAACTGGAAG	1353
QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1354	TGGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGTGTCATCTCTGTCATACTTT	1413
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1414	GTGCCCATGCACCTGTTTGGCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGTAT	1473
QY	381	GluLeThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	1474	GAGATCACGGAGGACACAGCCAGCAGTTGGCTGTGCCAACCGACGTCTCCCTATATACCC	1533

RESULT 9

US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55

; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores: 1.24e-204 Length: 8355
Pred. No.: 2114.00 Matches: 395
Score: 98.75% Conservative: 0

Best Local Similarity: 98.75%				Mismatches: 5	
Query Match: 99.16%				Indels: 0	
DB: 13				Gaps: 0	
SEQ14-X-AT-28-64-76 (1-400) x US-10-383-201-55 (1-8355)					
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20		
Db	1	ATGGACGTGAAGGAGAGAACCTTACCGCTCGCTGACCCGGCGCGACGCGGAGCGC	60		
QY	21	ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer	40		
Db	61	CGCTACACCAGCTCGTCCGCGGACAGCGAGGAGGGCAAAGCCCCGAGAAATCGTACAGC	120		
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60		
Db	121	TCCAGCGAGACCTGAAGGCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGCGTC	180		
QY	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80		
Db	181	AAGGACATTGTGCCGAGGCGGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG	240		
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100		
Db	241	CGGGAGCTGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCTGTACCGACAGACATT	300		
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120		
Db	301	GGCCTCCCCCACTGCGGCTACTCCATGGGGGTGGCTCTGATGCCGACATGGAGGCTGAC	360		
QY	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140		
Db	361	ACGGTGTCTGCCCTGAGCACCCCGTGCCTGTGTGGGGCCGGAGCACACGGTCAGGGCGC	420		
QY	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160		
Db	421	AGTCTCTGCCTGTCCAGCGGGCCAAATTCCAATCTCACACTCACCGACACCGAGCATGAA	480		
QY	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180		
Db	481	AACACTGAGACTGATCATCCGGGCGGCTGCAGAACCCAGCGGGCTCCGGACGCCGCGC	540		
QY	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200		
Db	541	CCGCCGCTCTCGACGCGCCACACCCCAACCCAGCACCCAGCGGCTCCATTAACTCCCTG	600		
QY	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220		
Db	601	AACCGGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGGCGGCGGCGGCGGCTCGCTC	660		
QY	221	SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240		
Db	661	TCCGGAGAGCCCCCTGCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTGGCTG	720		
QY	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr	260		
Db	721	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA	780		
QY	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280		
Db	781	TTGCAGGACAACCTCATTTAGATGGACATTCTCGGCGCTCCCGCCATGATGGGGCTTAC	840		
QY	281	SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer	300		
Db	841	AGTGACGGGCACTTCTCTTCAAGCCTGGAGGCACTCCCCGCTCTTCTGCACCAATCA	900		
QY	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320		
Db	901	CCAGGGTACCCACTGACGTCCAGCACAGTACTCTCTCCGCCCCGACCCCTGCCCGCGC	960		
QY	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340		
Db	961	AGCACCTTCGCCTGGCCGGCCCTTTAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG	1020		

QY	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360
Db	1021	TGGCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGGCATACTTT	1080
QY	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1081	GTGGCCATGACACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGCAGATGAT	1140
QY	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400
Db	1141	GAGATCACGAGGAGACACAGCCAGCAGTGGCCTGTGCCAACCGACGTCCTCCCTATACCCC	1200
RESULT 10			
US-10-042-865-1			
; Sequence 1, Application US/10042865			
; Publication No. US20040029216A1			
; GENERAL INFORMATION:			
; APPLICANT: Padigaru, Muralidhara			
; APPLICANT: Li, Li			
; APPLICANT: Zerhusen, Bryan D			
; APPLICANT: Casman, Stacie J			
; APPLICANT: Shenoy, Suresh G			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Zhong, Mei			
; APPLICANT: Gangolli, Esha A			
; APPLICANT: Burgess, Catherine E			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Vernet, Corine A.M			
; APPLICANT: Taylor, Sarah			
; APPLICANT: Tchernev, Velizar T			
; APPLICANT: Miller, Charles E			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Boldog, Ference L			
; APPLICANT: Grosse, William M			
; APPLICANT: Alsobrook II, John P			
; APPLICANT: Gerlach, Valerie L			
; APPLICANT: Edinger, Shlomit R			
; APPLICANT: Rothenberg, Mark E			
; APPLICANT: Ellerman, Karen			
; APPLICANT: MacDougall, John			
; APPLICANT: Malyankar, Uriel M			
; APPLICANT: Millet, Isabelle			
; APPLICANT: Peyman, John			
; APPLICANT: Smithson, Glennda			
; APPLICANT: Gunther, Erik			
; APPLICANT: Stone, David			
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of			
; TITLE OF INVENTION: Using the Same			
; FILE REFERENCE: 21402-537			
; CURRENT APPLICATION NUMBER: US/10/042,865			
; CURRENT FILING DATE: 2002-05-17			
; PRIOR APPLICATION NUMBER: 60/260,417			
; PRIOR FILING DATE: 2001-01-09			
; PRIOR APPLICATION NUMBER: 60/260,831			
; PRIOR FILING DATE: 2001-01-10			
; PRIOR APPLICATION NUMBER: 60/272,338			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: 60/274,876			
; PRIOR FILING DATE: 2001-03-09			
; PRIOR APPLICATION NUMBER: 60/284,704			
; PRIOR FILING DATE: 2001-04-18			
; NUMBER OF SEQ ID NOS: 264			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 8438			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-042-865-1			
Alignment Scores:			
Pred. No.:		2.31e-201	8438
Score:		2082.00	395

Percent Similarity: 98.01% Conservative: 0
Best Local Similarity: 98.01% Mismatches: 5
Query Match: 97.65% Indels: 3
DB: 13 Gaps: 1

SEQ14-X-AT-28-64-76 (1-400) x US-10-042-865-1 (1-8438)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 4 ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCGCGCGCGGACGCGAGCGC 63
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 64 CGCTACACCAGCTCGTCCGGGACAGCGAGGAGGGGCAAGCCCGCAGAAATCGTACAGC 123
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 124 TCCAGCGAGACCTGAAGCCTACGACCGCGCCCGCTAGCCTATGGCAGCGCGCTC 183
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 184 AAGGACATTGTGCCGCGAGGAGCGCGAGGAATTCTGCCGACAGGTGCCAACTTTCACCTG 243
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 244 CGGAGCTGGGCTGGAGAAGTAACGCCCCCTCACGGACCTGTACCGGACAGACATT 303
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 304 GGCTTCCCCACTGCGGCTACTCATGGGGCTGGCTCTGATCCGACATGGAGGCTGAC 363
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
Db 364 ACGTGCTGTCCCCTGAGCACCCCGTGCTGTGGGGCGGAGCACACGGTCAGGGCGC 423
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 424 AGCTCTGCCTGTCCAGCGGGGCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 483
QY 161 AsnThrGluThrAspHisProGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 484 AACACTGAGACTGATCATCCGGGGCGCTGCAGAACCAACCGCGGCTCCGGACGCGCGC 543
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 544 CCGCGCTCTCGCACGCCCCACACCCCAACACAGCACCAACCGCGGCTCCATTAACTCCCTG 603
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 604 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGCGCCCGCCAGCCACTCGCTC 663
QY 221 SerGlyGluProProAlaGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 664 TCCGAGAGCCCCCTGCCGGCGCGCCAGGAGCCTGCCACCGCCAGGAGAACTGGCTG 723
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 724 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGACA 783
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 784 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGCTTAC 843
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 844 AGTGACGGGCACTTCTCTTCAAGCCTTGAGGACCTCCCGCTCTTCTGCACCACATCA 903
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 904 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGCCAGCCCTGCCCGC 963
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340

Db 964 AGCACCTTGGCCCTGGCCGGCCTTTAACTCBAAGCCCTCCAAGTACTGTAACTGGAAG 1023
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db 1024 TGGCAGAGCCCTGAGCGCCATCGTATCTAGCCACTCTGGTCTATCCTGCTGGCATACTT 1083
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMet--- 379
Db 1084 GTGGCCATGCACCTGTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGACAGTACG 1143
QY 380 ---TyrGluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLe 398
Db 1144 GATTATGAGATCACGGAGGACACAGCCAGCAGTGTGGCTGTGCCAACCGACGTCCTCCT 1203
QY 398 uTyrPro 400
Db 1204 ATACCCC 1210
RESULT 11
US-10-290-578-11
; Sequence 11, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; Sliwkowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-No. US20030078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-290-578-11

Alignment Scores:
Pred. No.: 8.31e-121 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.27% Indels: 0
DB: 15 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-10-290-578-11 (1-2387)

QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
| | | | |
Db 1 CATCCGGCGGCTGCAGAACCCACGCGGGCTCCGGACGCCGCGCGCTCTCGCAC 60
| | | | |
QY 186 AlaHisThrProAsnGlnHisHisAlaAaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
| | | | |
Db 61 GCCCACACCCCAACCAGCACCCACGCGGCTCCATTAACTCCCTGAACCGGGCAACTC 120
| | | | |
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
| | | | |
Db 121 ACGCCGAGGAGCAACCCAGCCCGGCCCCCAGGACCACTCGCTCTCCGAGAGCCCCCT 180
| | | | |
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
| | | | |
Db 181 GCCGGCGGCGCCAGGAGCCTGCCCCAGCCCGGAGAACTGGTGCTCAACAGCAACATC 240
| | | | |
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265
| | | | |
Db 241 CCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGACATTGCAGGACACCTC 300
| | | | |
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
| | | | |
Db 301 ATTGAGATGGACATTCTCGCGCTCCCGCCATGATGGGCTTACAGTGACGGGCACTTC 360
| | | | |
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
| | | | |
Db 361 CTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCATCACAGGGTACCCACTG 420
| | | | |
QY 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325
| | | | |
Db 421 ACGTCCAGCACAGTGTACTCTCTCCGCCCCGACCCCTGCCCCGAGCACCTTCGCCCGG 480
| | | | |
QY 326 ProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345
| | | | |
Db 481 CCGGCCTTTAACTCAAGAGCCCTCCAAGTACTGTAACTGGAAGTGCGCAGCCCTGAGC 540
| | | | |
QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365
| | | | |
Db 541 GCCATCGTCATCTCAGCCACTCTGGTCATCTCTGTGSCATACTTTGTGGCCATGCACCTG 600
| | | | |
QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385
| | | | |
Db 601 TTTGGCCTAACTGGCACCTGCAGCCGATGGAGGGCAGATGTATGAGATCAGGAGGAC 660
| | | | |
QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
| | | | |
Db 661 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATACCCC 705
| | | | |

RESULT 12

US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1

GENERAL INFORMATION:

; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca

; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37

; LENGTH: 8645

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-038-854-37

Alignment Scores:

Pred. No.: 1.92e-88 Length: 8645
Score: 975.00 Matches: 210
Percent Similarity: 62.41% Conservative: 49
Best Local Similarity: 50.60% Mismatches: 92
Query Match: 45.73% Indels: 64
DB: 17 Gaps: 12

SEQ14-X-AT-28-64-76 (1-400) x US-10-038-854-37 (1-8645)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
| | | | |
Db 151 ATGGATGTGAAAGAACGCGGCCCTTACTGCTCCCTGACCAAGAGCAGACGAGAGGAA 210
| | | | |
QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
| | | | |
Db 211 CGGCGCTACAAATTCCTCCGAGACAAATGAGGAGTGCCGGGTACCCACACAGAAGTCC 270
| | | | |
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
| | | | |
Db 271 TACAGTTCAGCGAGACATTGAAAGCTTTTGATCATGATTCTCGCGGCTGCTTTACGGC 330
| | | | |
QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
| | | | |
Db 331 AACAGAGTGAAGGATTTGGTTTCACAGAGAAGCAGACGAGTTCTACTAGACAAGGACAGAT 390
| | | | |
QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
| | | | |
Db 391 TTTACCTTAAGGCAGTTAGGAGTTGTGAACCACTCGAAGAGGACTGGCATTTTGT 450
| | | | |
QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
| | | | |


```

;
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderne, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catharine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
;
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
;
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
;
US-10-038-854-35

```

Alignment Scores:

Pred. No.:	1.93e-88	Length:	8675
Score:	975.00	Matches:	210
Percent Similarity:	62.41%	Conservative:	49
Best Local Similarity:	50.60%	Mismatches:	92
Query Match:	45.73%	Indels:	64
DB:	17	Gaps:	12

SEQ14-X-AT-28-64-76 (1-400) x US-10-038-854-35 (1-8675)

QY 1 MetAspVallysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19

Db 151 ATGGATGTGAAAGAACGCGAGCCCTTACTGCTCCCTGACCAAGAGACGAGAGAGAA 210

QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38

Db 211 CGCGGCTACACAAATTCCTCCGACAGACATGAGGAGTCCCGGGTACCCACACAGAAAGTCC 270

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
Db 271 TACAGTTCAGCGAGACATTGAAAGCTTTTGATCATGATTCCTCGGGCTGCTTTACGGC 330

QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77
Db 331 AACAGAGTGAAGGATTTGGTTTCACAGAGAAGCAGACGAGTTCACAGCAAGGACAGAA 390

QY 78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
Db 391 TTTACCTTAAGGCAGTTAGGAGTTTGTGTAACCACTCGAAGAGGACTGGCATTTTGT 450

QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
Db 451 GCGGAAATGGGGCTCCCTCACAGAGGTTACTCTATCAGTGCAGGGTCAGATGCTGATACT 510

QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
Db 511 GAAATGAAGCAGTGATGTCCCCAGAGCATGCCATGAGACTTTGGGCGAGGGGTCAAA 570

QY 138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
Db 571 TCAGGCCGCGAGCTCCTGCCTGTCAAGTCGGTCCAACCTCAGCCCTCACCTGACAGATACG 630

QY 158 GluHisGluAsn-----ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAla 174
Db 631 GAGCAGAAACAAGTCCGACAGTGAGATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 690

QY 175 ArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHisHisAla 194
Db 691 ACCCTGCAGCCCTTGCCGCCT-----TCCCATAGCAGCACTCTGCACAGCATCAT--- 741

QY 195 AlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAla 214
Db 742 CCATCCATCACTTCTCAACAGAAACTCCCTGACCAATAGAAAGAACCCAGAGTCCGGCC 801

QY 215 ProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGln----- 230
Db 802 CCG-----CCGGCTGCTTTGCCCGCCGAGCTGCAAAACCACA 837

QY 231 ---GluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThr 249
Db 838 CCCGAGTCCGTCCAGCTGCAGGACAGCTGGGTCTTGGCAGTAATGTACCACTGGAAAGC 897

QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 898 AGG----- 900

QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 901 -----CATTTCTCTATTCAAAACA 918

QY 290 Gly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSer 308
Db 919 GGAACAGGTACAAAGCCACTGTTTCAGTACTGCAACCCCGAGGATACACAATGGCATCTGGC 978

QY 309 ThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArgProAlaPhe 328
Db 979 TCTGTTTATTACCACTTACTCGGCCACTACCTAGAAACACCCCTATCAGAGAGTGTCTTT 1038

QY 329 AsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleVal 348
Db 1039 AAATTCAAGAAAGTCTTCAAAGTACTGTAGCTGGAAATGCACCTGCTGTGCGGTAGGG 1098

QY 349 IleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeu 368
Db 1099 GTCTCGGTGCTCTGGCAATACCTCCTGTCTTATTTATAGCAATGCATCTCTTTGGCCTC 1158

QY 369 AsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGlu-----IleThrGluAsp 385
Db 1159 AACTGGCAGCTACAGCAGACTGAAAAATGACACATTTGAGAATGGAAAGTGAATTCTGAT 1218

QY 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1219 ACC-----ATGCCAACAAACACTGTGTCTATTACCT 1248

RESULT 14
US-10-144-194A-81
; Sequence 81, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: Origene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 9695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (435)..(8336)
US-10-144-194A-81

Alignment Scores:
Pred. No.: 1.68e-84 Length: 9695
Score: 937.00 Matches: 206
Percent Similarity: 63.57% Conservative: 54
Best Local Similarity: 50.37% Mismatches: 93
Query Match: 43.95% Indels: 56
DB: 16 Gaps: 13

SEQ14-X-AT-28-64-76 (1-400) x US-10-144-194A-81 (1-9695)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
Db 435 ATGGATGTAAAGGACCGCGA---CACCGCTCTTTGACCAGAGGACGCTGTGGCAAGAG 491

QY 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 492 TGTCGCTACACAGCTCTCTCTGGACAGTGAGGACTGCCGCTGCCACACAGAAATCC 551

QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 552 TACAGCTCCAGTGAGACTCTGAAGGCCTATGACCATGACAGCAGGATGCATATGGAAC 611

QY 59 ArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPhe 78
Db 612 CGAGTCACAGACTCATCCACCGGAGTGATGATGATTTCTTAGACAAGAACCAACTTC 671

QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThr 98
Db 672 ACCCTTGCCGAAGTGGGCATCTGTGAGCCCTCC---CCACACCGAAGCGGCTACTGCTCC 728

QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 729 GACATGGGATCCTTCACAGGGCTACTCCCTTAGCACAGGGTCTGACGCGGACTCCGAC 788

QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
Db 789 ACCGAGGAGGATGTCTCCAGAACACGCCATCAGACTGTGGGGCAGAGGATATAAATCC 848

QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 849 AGGCGCAGTTCGGCCTGTCCAGTCGTGAAAACTCGGCCCTTACCTGACTGACTCTGAC 908

QY 159 HisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArg-----Leu 176
Db 909 AACGAAACAAATCAGATGATGAGAACGGCCCTCCGAACCCACACGCGGACTCGACTCTG 968

QY 177 ArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaSer 196
Db 969 AGGCCCCCTCTCCCAACC---CCTCACAAACACACAGCTGTCCCATCACAC---TCGTCC 1022

```
QY 197 IleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsn-----ProSer 212
Db 1023 GCCAACTCCCTCAACAGGAACTCACTGACCACATCGCGGAGTCAGATCCACGCCCGGCC 1082
QY 213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro 232
Db 1083 CCAGCGCCCAATGAC-----CTGGCCACCACACCA-----GAGTCC 1118
QY 233 AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu 252
Db 1119 GTTCAGCTTCAGGACAGCTGGGTGCTAAACAGCAACGTGCCACTGGAGACCCGG----- 1172
QY 253 GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly 272
Db 1172 ----- 1172
QY 273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLys---ProGlyGly 291
Db 1173 -----CACTTCCTCTTCAAGACCTCCTCGGGG 1199
QY 292 ThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyr 311
Db 1200 AGCACACCCCTTGTTTCAGCAGCTCTTCCCGGGATACCCCTTGACCTCAGGAACGGTTTAC 1259
QY 312 SerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeuLys 331
Db 1260 ACGCCCCCGCCCGCTGCTGCCCAGGAATACTTTCTCCAGGAAGGCTTTCAGAGCTGAAG 1319
QY 332 LysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAla 351
Db 1320 AAGCCCTCCAAATACTGCAGCTGGAATGTGCTGCCCTCTCCGCCATTGCCCGGCCCTC 1379
QY 352 ThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrpHis 371
Db 1380 CTCCTGGCTATTTTGTGGCGTATTTTCATAGCAATGCATCTGCTCGGACTCAATTGGCAA 1439
QY 372 LeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrpPro 391
Db 1440 CTCGAGCCTGCAGATGGGCACACCTTT-----AACATGGGATAAGGACCGGCTTACCA 1493
QY 392 ValProThrAspValSerLeuTyrPro 400
Db 1494 GGAAACGATGATGTGGCAACAATGCCA 1520

RESULT 15
US-09-808-602-75
; Sequence 75, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Heirman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 75
; LENGTH: 2496
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-808-602-75
```

```
Alignment Scores: 5.05e-82 Length: 2496
Pred. No.: 905.00 Matches: 208
Score: 54.28% Conservative: 52
Best Local Similarity: 43.42% Mismatches: 85
Query Match: 42.45% Indels: 134
DB: 9 Gaps: 14

SEQ14-X-AT-28-64-76 (1-400) x US-09-808-602-75 (1-2496)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArg---AspAlaGlu 19
Db 1 ATGGATATAAAAGATCGAAGA---CACCGCTCTTTGACGAGAGGCCGGTTCGGGAAGGAG 57
QY 20 ArgArgTyrThrSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38
Db 58 TGTGCTATATACTAGTTCTTCACTCGACAGTGAAGACTGCAGAGTACCAACTCAGAAGTCC 117
QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer 58
Db 118 TACAGTCCAGTGAGACTCTGAAGCATATATGACCATGACACGAGGATGCACGAAAT 177
QY 59 ArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPhe 78
Db 178 CGAGTTTCAGACCTGGTTCACAGGGAGTCCGATGATGTTTCCAAGGCAAGAACGAACTTC 237
QY 79 ThrLeuArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThr 98
Db 238 ACCCTTGAGAACTGGGAATCTGTGAGCCCTCT---CCCCATCGAAGTGGCTACTGCTCG 294
QY 99 AspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu 118
Db 295 GACATAGGAATACTCCATCAAGGCTATTCTTTGAGCACTGGCTCTGATGCTGACTCAGAC 354
QY 119 AlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSer 138
Db 355 ACGGAGGGCGGATGTCTCCAGAGCACGCGATCAGGCTGTGGGGAAGAGGGATCAAAATCC 414
QY 139 GlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGlu 158
Db 415 AGCGGAAGTTCTGGCCTGTCAAGTCGTGAAAACTCGGCTCTCACGCTCACTGACTCCGAC 474
QY 159 HisGluAsnThrGluThrAspHisProGly----- 168
Db 475 AATGAGAACAAAGTCAGATGAGGAAAACGGTCGTCCTCCATTCACCTACATCCTCGTCTAGC 534
QY 168 ----- 168
Db 535 CTCTCTCCCATCTGCTCAGCTGCCAGTTCTCATAATCTCTCCACCAGTTAGCTGCCAGATG 594
QY 168 ----- 168
Db 595 CCATTGCTAGACAGCAATACGTCCCATCAAATCATGGACACCAATCCTGACGAGGAGTTTC 654
QY 169 -----Gly 169
Db 655 TCTCTAAATTATACCTACTAAGAGCATGTTTCAGGGCCACAGCAGGCATCCAGCAGTGGC 714
QY 170 LeuGlnAsnHis-----AlaArgLeuArgThrPro---ProProProLeuSerHis 185
Db 715 CCTTCAAACCATCACAGCCAGTCAACGCTGAGGCCACCTCTCCCCCTCTCTCACAAACCAC 774
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 775 TCGCTGTCC-----CATCATCACTCGTCTGCCAACTCCCTCAACAGGAACCTCGCTC 825
QY 206 ThrProArgSerAsn-----ProSerProAlaProThrAspHisSerLeuSer 221
Db 826 ACCAACCGCCGCAACCAGATCCACGGCCTGCTCCCGCTCCCAATGAC-----CTGGCG 879
QY 222 GlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeu 241
```



```
Db      880 ACCACGCCT-----GAGTCTGTGCAGCTGCAGGACAGCTGGGTGCTC 921
QY      242 AsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThrLeu 261
Db      922 AACAGCAACGTGCCGCTGGAGACCAGG----- 948
QY      262 GlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSer 281
Db      948 ----- 948
QY      282 AspGlyHisPheLeuPheLys---ProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db      949 -----CATTCTTGTTTAAACATCTTCTTGGAACGACTCCGCTGTTCACTAGCTCTTCC 1002
QY      301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db      1003 CTGGCTACCCACTGACCTCAGGAACAGTTTATATACTCCACCTCCAGGCTGTTACCTAGA 1062
QY      321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db      1063 AATACATTTTCCAGGAATGCATTCAAGCTGAAAAGCCCTCCAAGTATTGTAGCTGGA 1122
QY      341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
Db      1123 TGTGCTGCTTTATCTGCAATTGCTGCTGCAGTCCCTGCTTGCCATCCTGCTAGCATATTTC 1182
QY      361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db      1183 ATAGCGATGCACCTCCTGGGGCTGAACTGGCAGCTGCAGCCCGCGGACGGACACCTTC 1242
QY      381 GluIle-----ThrGluAspThrAlaSerSerTrpPro 391
Db      1243 AGCAACGGGCTGGGGCGGGCGGGCGGGCGGAGGACGGAGCGGGCGGCCACCT 1299
```

Search completed: August 15, 2004, 00:56:30
Job time : 659 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 19:26:15 ; Search time 103 Seconds
(without alignments)
2155.149 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76
Perfect score: 2132
Sequence: 1 MDVKERKPYRSLTRRRDAER.....EITEDTASGWPVPTDVS LYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues 1365418
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/MITRA020/runat_06082004_114102_237/app_query.fasta.1.583
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MITRA020 @CGN 1 1 69 @runat_06082004_114102_237 -NCPU=6 -ICPU=3 -NO_MMAP
-LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2122	99.5	1680	3 US-08-891-845-3	Sequence 3, Appli
2	2122	99.5	1680	4 US-09-514-573-3	Sequence 3, Appli
3	2122	99.5	3111	3 US-08-891-845-1	Sequence 1, Appli
4	2122	99.5	3111	4 US-09-514-573-1	Sequence 1, Appli
5	1285	60.3	2387	3 US-08-891-845-11	Sequence 11, Appl
6	1285	60.3	2387	4 US-09-514-573-11	Sequence 11, Appl
7	131.5	6.2	2949	4 US-09-623-326-6	Sequence 6, Appli
8	131.5	6.2	4446	4 US-09-231-899-69	Sequence 69, Appl
C 9	129.5	6.1	1860	4 US-09-051-994-1	Sequence 1, Appli
C 10	128.5	6.0	8438	1 US-07-945-283-1	Sequence 1, Appli
11	127.5	6.0	4456	4 US-09-095-443-1	Sequence 1, Appli
C 12	126.5	5.9	4403765	3 US-09-103-840A-2	Sequence 2, Appli

C 13	126.5	5.9	4411529	3 US-09-103-840A-1	Sequence 1, Appli
14	126	5.9	2507	2 US-08-471-066B-1	Sequence 1, Appli
15	125	5.9	1998	4 US-09-252-991A-11748	Sequence 11748, A
16	125	5.9	3180	4 US-09-252-991A-11618	Sequence 11618, A
C 17	123.5	5.8	2499	4 US-09-758-282B-100	Sequence 100, App
18	123.5	5.8	2517	4 US-09-758-282B-147	Sequence 147, App
19	123.5	5.8	2517	4 US-09-758-282B-149	Sequence 149, App
20	123.5	5.8	2517	4 US-09-758-282B-152	Sequence 152, App
21	123.5	5.8	2517	4 US-09-758-282B-168	Sequence 168, App
22	123.5	5.8	2517	4 US-09-758-282B-171	Sequence 171, App
23	123.5	5.8	2517	4 US-09-758-282B-180	Sequence 180, App
24	123.5	5.8	2517	4 US-09-758-282B-183	Sequence 183, App
25	123.5	5.8	2517	4 US-09-758-282B-186	Sequence 186, App
26	123.5	5.8	2733	4 US-09-623-326-1	Sequence 1, Appli
27	123	5.8	4078	4 US-09-016-434-1152	Sequence 1152, Ap
28	123	5.8	4553	4 US-09-023-655-975	Sequence 975, App
29	123	5.8	4689	3 US-09-105-537-34	Sequence 34, Appl
30	123	5.8	5121	4 US-09-252-991A-15189	Sequence 15189, A
31	123	5.8	36778	3 US-09-105-537-5	Sequence 5, Appli
32	123	5.8	38506	3 US-09-320-878-19	Sequence 19, Appl
33	123	5.8	38506	4 US-09-141-908-1	Sequence 1, Appli
34	123	5.8	38506	4 US-09-657-440-19	Sequence 19, Appl
C 35	123	5.8	71989	4 US-09-443-501A-2	Sequence 2, Appli
36	122.5	5.7	1686	2 US-08-648-657-1	Sequence 1, Appli
37	122.5	5.7	1689	2 US-08-648-657-2	Sequence 2, Appli
38	122.5	5.7	2517	4 US-09-758-282B-174	Sequence 174, App
39	122.5	5.7	2588	4 US-09-480-017-3	Sequence 3, Appli
40	122.5	5.7	2943	4 US-09-379-530B-3	Sequence 3, Appli
41	122.5	5.7	49377	1 US-08-764-233A-1	Sequence 1, Appli
42	121.5	5.7	1682	3 US-09-096-399-1	Sequence 1, Appli
43	121.5	5.7	1696	3 US-09-096-399-3	Sequence 3, Appli
44	121.5	5.7	1899	4 US-09-640-958-3	Sequence 3, Appli
45	121.5	5.7	1965	4 US-09-640-958-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-891-845-3
; Sequence 3, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-891-845-3

Alignment Scores:

Pred. No.: 2.05e-188 Length: 1680
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 3 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-08-891-845-3 (1-1680)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 1 ATGGACGTGAAGGAGAGGAGCCCTTACCGCTCGTGACCCGGCGCGACGCGGAGCGC 60

QY 21 ArgTyrThrSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 61 CGCTACACCAAGTCGTCCGGGACAGCGAGGAGGCAAGCCCGCAGAAATCGTACAGC 120

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 121 TCCAGCGAGACCCCTGAAGCCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGCGTC 180

QY 61 LysAspIle***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
Db 181 AAGGACATTGTGCGCGCAGGAGCGCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCGTG 240

QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 241 CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGACCCCTGTACCGGACAGACATT 300

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 301 GGCCTCCCACTGCGGCTACTCCATGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 360

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 361 ACGGTGCTGTCCCTGAGCACCCCGTGTGCTGTGGGGCGGAGCACACGGTCAGGGCGC 420

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 421 AGCTCCTGCCTGTCCAGCCGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 480

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 481 AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCCAGCGCGGCTCCGGACGCCGCG 540

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
Db 541 CCGCCGCTCTCGACGCCCCACACCCCAACACGACCAACCGCGGCTCCATTAACTCCCTG 600

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 601 AACCGGGGCAACTTCACGCCGAGGAGCAACCCCGAGCCCGCCCGCCCGGACCACTCGCTC 660

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 661 TCCGGAGAGCCCTGTCCGGCGCGCCAGGAGCCTGCCACCGCCAGGAGAACTGGCTG 720

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
Db 721 CTCAACAGCAACATCCCTCTGGAGACCAGAAACCTAGGCAAGCAGCATTCCTAGGGACA 780

QY 261 LeuGlnAspAsnIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 781 TTGCAGGACAACTCATTTGATGGACATCTCGGGCGCCTCCCGCCCATGATGGGGCTTAC 840

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
Db 841 AGTGACGGGCACTTCTCTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCATCA 900

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuProArg 320
Db 901 CCAGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCCGC 960

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysProSerLysTyrCysAsnTrpLys 340
Db 961 AGCACCTTCGCCCGCGCGCTTTAACCTCAAGAGCCCTCCAAGTACTGTAAGTGAAG 1020

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1021 TGCGCAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCACTCTGCTGGCATACTTT 1080

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1081 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1140

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1141 GAGATACGAGGAGACACAGCCAGCAGTGTGGCCTGTGTGCCAACCGACGTCTCCCTATACCCC 1200

RESULT 2

US-09-514-573-3
; Sequence 3, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-3

Alignment Scores: 2.05e-188 Length: 1680
Pred. No.:

Score:		2122.00	Matches:	396
Percent Similarity:		99.00%	Conservative:	0
Best Local Similarity:		99.00%	Mismatches:	4
Query Match:		99.53%	Indels:	0
DB:		4	Gaps:	0
SEQ14-X-AT-28-64-76 (1-400) x US-09-514-573-3 (1-1680)				
Qy	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20	
Db	1	ATGGACGTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCGGACGCCGAGCGC	60	
Qy	21	ArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyrSer	40	
Db	61	CGCTACACAGACTCGTCCGCGGACACGAGGAGGGCAAAGCCCCCGAGAAATCGTACAGC	120	
Qy	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60	
Db	121	TCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTC	180	
Qy	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80	
Db	181	AAGGACATGTGCCGCGCAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCTG	240	
Qy	81	ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle	100	
Db	241	CGGAGTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT	300	
Qy	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120	
Db	301	GGCTTCCCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	360	
Qy	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140	
Db	361	ACGGTGCTGTCCCTGAGCACCCCGTGGTCTGTGGGGCCGGAGCACACGGTCAGGGCGC	420	
Qy	141	SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu	160	
Db	421	AGCTCCTGCTGTCCAGCCGGGCCAATTCCAATCTCACACTACCGACACCGAGCATGAA	480	
Qy	161	AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro	180	
Db	481	AACACTGAGACTGATCATCCGGGGCGGCTGCAGAACCCACGGCGGCTCCGGACGCCGCG	540	
Qy	181	ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu	200	
Db	541	CCGCGGCTCTGCACGCCCCACACCCCCAACCCAGCACCCACGGCGGCTCCATTACTCCCTG	600	
Qy	201	AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu	220	
Db	601	AACGGGGCAACTTCACGCCGAGGAGCAACCCACGCCCGCCCGCCACGGACCACTCGCTC	660	
Qy	221	SerGlyGluProProAlaGlyGlyValaGlnGluProAlaHisAlaGlnGluAsnTrpLeu	240	
Db	661	TCCGAGAGCCCCCTGCGGGCGGCGCCAGGAGCCTGCCACGCCCCAGGAACTGGCTG	720	
Qy	241	LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr	260	
Db	721	CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATTCTTAGGGACA	780	
Qy	261	LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr	280	
Db	781	TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGGCTTAC	840	
Qy	281	SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer	300	
Db	841	AGTGACGGCACTTCTCTTCAAGCTGGAGGCACCTCCCGGCTCTTCTGTACCCACATCA	900	
Qy	301	ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg	320	
Db	901	CCAGGTACCCCACTGACGTCCAGCACAGTGTACTCTCCTCCGCCCCGACCCCTGCCCGCG	960	
Qy	321	SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys	340	

Db	961	AGCACCTTCGCCCGCGCGCCTTTAACTCAAGAGCCCTCCAAGTACTGTAACCTGGAAG	1020	
Qy	341	CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe	360	
Db	1021	TGCGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCCTCTGGCATACTTT	1080	
Qy	361	ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380	
Db	1081	GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT	1140	
Qy	381	GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro	400	
Db	1141	GAGATCACGAGGACACAGCCAGCAGTGGCCCTGTGCCAACCCGACGTCCTCCCTATACCCC	1200	
RESULT 3				
US-08-891-845-1				
; Sequence 1, Application US/08891845				
; Patent No. 6096873				
; GENERAL INFORMATION:				
; APPLICANT: Schaefer, Gabriele M.				
; APPLICANT: Sliwowski, Mark				
; TITLE OF INVENTION: Gamma-Heregulin				
; NUMBER OF SEQUENCES: 11				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Genentech, Inc.				
; STREET: 460 Point San Bruno Blvd				
; CITY: South San Francisco				
; STATE: California				
; COUNTRY: USA				
; ZIP: 94080				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk				
; COMPUTER: IBM PC compatible				
; OPERATING SYSTEM: PC-DOS/MS-DOS				
; SOFTWARE: WinPatin (Genentech)				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/891,845				
; FILING DATE:				
; CLASSIFICATION: 435				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: 60/021640				
; FILING DATE: 07/12/96				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Lee, Wendy M.				
; REGISTRATION NUMBER: 40,378				
; REFERENCE/DOCKET NUMBER: P1043				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 415/225-1994				
; TELEFAX: 415/952-9881				
; TELEX: 910/371-7168				
; INFORMATION FOR SEQ ID NO: 1:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 3111 base pairs				
; TYPE: Nucleic Acid				
; STRANDEDNESS: Single				
; TOPOLOGY: Linear				
US-08-891-845-1				
Alignment Scores:				
Pred. No.:		5.43e-188	Length:	3111
Score:		2122.00	Matches:	396
Percent Similarity:		99.00%	Conservative:	0
Best Local Similarity:		99.00%	Mismatches:	4
Query Match:		99.53%	Indels:	0
DB:		3	Gaps:	0
SEQ14-X-AT-28-64-76 (1-400) x US-08-891-845-1 (1-3111)				
Qy	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20	
Db	334	ATGACGCTGAAGGAGAGGAAGCCTTACCGCTCGCTGACCCGGCGCCGACCGAGCGC	393	

QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
|||
Db 394 CGCTACACAGCTCGTCCGCGGACAGCGAGGGCAAAGCCCCGAGAAATCGTACAGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||
Db 454 TCCAGCGAGACCCCTGAAGGCCTACGACCAGGACGCCCGCTAGCTATGGCAGCCGCGTC 513

QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
|||
Db 514 AAGGACATTGTGCCGCGCAGGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCGTG 573

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
|||
Db 574 CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCCCTGTACCGGACAGACATT 633

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
|||
Db 634 GGCCTCCCCCACTGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC 693

QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
|||
Db 694 ACGGTGCTGTCCCTGAGCACCCCGTGCCTGTGTGGGGCCGAGACACGGTCAGGGCGC 753

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
|||
Db 754 AGTCTCTGCTGTCCAGCGGGGCCAATTCCAATCTCACACTCACCGACACCGAGCATGAA 813

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
|||
Db 814 AACACTGAGACTGATCATCCGGCGGCTTCGAGAACCACGCGGGCTCCGGACGCCGCGG 873

QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeu 200
|||
Db 874 CCGCGCTCTCGACGCGCCACACCCCAACAGCACCAACGCGGCTCCATTAACTCCCTG 933

QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
|||
Db 934 AACCGGGCAACTTCACGCGGAGGAGCAACCCAGCCCGGCCCAACGACCACTCGCTC 993

QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
|||
Db 994 TCCGGAGAGCCCTGCCGGCGGCCAGGAGCCTGCCACGCCCAAGAGAACTGGCTG 1053

QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThr 260
|||
Db 1054 CTCACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCCATCTCTAGGGACA 1113

QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
|||
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGGCGCTCCCGCCATGATGGGGCTTAC 1173

QY 281 SerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSer 300
|||
Db 1174 AGTACGGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGACCAATCA 1233

QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
|||
Db 1234 CCAGGGTACCCACTGACGTCCAGCACAGTGAATCTCCCGCCCGACCCCTGCCCGC 1293

QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
|||
Db 1294 AGCACCTTCGCCCCGGCGCCTTTAACCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAG 1353

QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPhe 360
|||
Db 1354 TGGCAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCATCTCTGCTGGCATACTTT 1413

QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
|||
Db 1414 GTGGCCATGCACCTGTTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGCAGATGTAT 1473

QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 1474 GAGATCAGGAGGACACAGCCAGCAGTTGGCTGTGCCAACCGACGTCCTCCCTATACCCC 1533

RESULT 4
US-09-514-573-1
; Sequence 1, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-09-514-573-1

Alignment Scores:
Pred. No.: 5.43e-188 Length: 3111
Score: 2122.00 Matches: 396
Percent Similarity: 99.00% Conservative: 0
Best Local Similarity: 99.00% Mismatches: 4
Query Match: 99.53% Indels: 0
DB: 4 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-09-514-573-1 (1-3111)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
|||
Db 334 ATGGACGTGAGGAGAGGAAAGCCTTACCCTCGCTGACCGCGCGCGACCGCGCGC 393

QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
|||
Db 394 CGCTACACCACTCGTCCCGCGGACAGCGAGGGCAAAGCCCCGAGAAATCGTACAGC 453

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||
Db 454 TCCAGCGAGACCTGAAGGCCTACGACCAGGACGCCCGCTAGCCTATGGCAGCCGCGTC 513

QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
|||
|||

Db 514 AAGGACATTGTCCGACGAGGCCGAGGAATTCTGCCGCACAGGTGCCAACTTCACCCCTG 573
QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 574 CGGAGCTGGGGCTGGAAGAAGTAACGCCCCCTCACGGGACCTGTACCGGACAGACATT 633
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 634 GGCCTCCCCCACTGGGGCTACTCCATGGGGCTGGCTCTGATGCCACATGGAGGCTGAC 693
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 694 ACGGTGCTGTCCCTGAGCACCCCGCTGCGTCTGTGGGGCCGGAGCACACGCTCAGGGCGC 753
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 754 AGCTCTCGCTGTCCAGCCGGGCCAATTCCAAATCTCACACTCACCGACACCGAGCATGAA 813
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 814 AACACTGAGACTGATCATCCGGCGGCTGCAGAACCAACCGCGGCTCCGAGCCGCCG 873
QY 181 ProProLeuSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeu 200
Db 874 CCGCGCTCTCGACGCCACACCCCCAACCCAGCACACCGCGGCTCCATTAACTCCCTG 933
QY 201 AsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeu 220
Db 934 AACCGGGCAACTTCAGCCGAGGAGGAGCAACCCAGCCCGGCCCCACGACCACTCGCTC 993
QY 221 SerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeu 240
Db 994 TCCGGAGAGCCCCCTGCGCGGCGCCAGAGCCTGCCACGCCCGAGAGAACTGGCTG 1053
QY 241 LeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThr 260
Db 1054 CTCAACAGCAACATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCATTCTTAGGGACA 1113
QY 261 LeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr 280
Db 1114 TTGCAGGACAACCTCATTTGAGATGGACATTCTCGGCGCTCCCGCCATGATGGGCTTAC 1173
QY 281 SerAspGlyHisPheLeuPheLysProGlyThrSerProLeuPheCysThrThrSer 300
Db 1174 AGTGACGGGCACTTCCTCTTCAAGCCTTGAGGACCTCCCGCTCTTGTGACCAATCA 1233
QY 301 ProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuProArg 320
Db 1234 CCAGGTACCCACTGACGTCCAGCACAGTGTAATCTCTCCGCCCGGACCCCTGCCCGCG 1293
QY 321 SerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLys 340
Db 1294 AGCACTTCGCCCCGGCCGCTTTAACTCAAGAACGCCCTCCAAGTACTGTAACTGGAAG 1353
QY 341 CysAlaAlaLeuSerAlaIleValIleSerAlaThrLeuValIleLeuAlaTyrPhe 360
Db 1354 TGCGAGCCCTGAGCGCCATCGTCATCTCAGCCACTCTGGTCACTCTGTGGCATACTTT 1413
QY 361 ValAlaMetHisLeuPheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380
Db 1414 GTGGCATGCACCTGTTTGGCTAACTGGCACCTGCAGCCGATGGAGGGGCAGATGAT 1473
QY 381 GluIleThrGluAspThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400
Db 1474 GAGATCAGGAGGACACAGCCAGCATGTGGCCTGTGCCAACCGACGTCTCCCTATACCC 1533

RESULT 5

US-08-891-845-11
; Sequence 11, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark

; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-891-845-11

Alignment Scores:

Pred. No.: 4.83e-110 Length: 2387
Score: 1285.00 Matches: 235
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.27% Indels: 0
DB: 3 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-08-891-845-11 (1-2387)

QY 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSerHis 185
Db 1 CATCCGGCGGCGCTGCAGAACCAACGCGCGGCTCCGGACGCCGCGCGCTCTCGCAC 60
QY 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205
Db 61 GCCACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAATC 120
QY 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225
Db 121 ACGCGAGGAGCAACCCAGCCCGGCCCGCCCGACCACTCGCTCTCCGGAGAGCCCTT 180
QY 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245
Db 181 GCCCGCGGCGCCAGGAGCCTGCCACGCCCGAGGAGAACTGGCTGCTCAACAGCAACATC 240
QY 246 ProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGlyThrLeuGlnAspAsnLeu 265
Db 241 CCCTGGAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTC 300
QY 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285
Db 301 ATGAGATGGACATTCTCGCGCCTCCCGCCATGATGGGCTTACAGTGACGGGCACTTC 360
QY 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305

Db 361 CTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCACCAGGGTACCCACTG 420

Qy 306 ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg 325

Db 421 ACGTCCAGCACAGTGTACTCTCTCGCCCCGACCCCTGCCCGCAGCACCTTCGCCCGG 480

Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345

Db 481 CCGGCTTTAACTCAAGAGCCCTCCAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540

Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365

Db 541 GCCATCGTCATCTCAGCCACTCTGGTCATCTGCTGGCATACTTTGTGGCCATGCACCTG 600

Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385

Db 601 TTTGGCCTAAACTGGCACCTGCAGCCGATGGAGGGGAGATGTATGAGATCACGGAGGAC 660

Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 661 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATATACCCC 705

RESULT 6

US-09-514-573-11

; Sequence 11, Application US/09514573

; Patent No. 6500941

; GENERAL INFORMATION:

; APPLICANT: Schaefer, Gabriele M.

; APPLICANT: Sliwkowski, Mark

; TITLE OF INVENTION: Gamma-Herregulin

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 460 Point San Bruno Blvd

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/514,573

FILING DATE: 28 FEB 2000

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/891845

FILING DATE: 10 JULY 1997

APPLICATION NUMBER: 60/021640

FILING DATE: 07/12/96

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378

REFERENCE/DOCKET NUMBER: P1043

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994

TELEFAX: 415/952-9881

TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 2387 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Single

TOPOLOGY: Linear

US-09-514-573-11

Alignment Scores:

Pred. No.: 4.83e-110 Length: 2387

Score: 1285.00 Matches: 235

Percent Similarity: 100.00% Conservativity: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 60.27% Indels: 0

DB: 4 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x US-09-514-573-11 (1-2387)

Qy 166 HisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeuSerHis 185

Db 1 CATCCGGCGGCTGCAGAACACCGCGGGTCCGGACGCGCGCGCGCTCTCGCAC 60

Qy 186 AlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPhe 205

Db 61 GCCACACCCCCAACACAGCACCCACGCGGCTCCATTAATCTCCTGAACCGGGGCACTTC 120

Qy 206 ThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProPro 225

Db 121 ACGCCGAGGAGCAACCCAGCCCGCGCCCCCAGCACCTCGCTCTCCGAGAGCCCCCT 180

Qy 226 AlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIle 245

Db 181 GCCGGCGCGCCAGGAGCCTGCCACGCGCCAGGAGAACTGGCTGCTCAACAGCAACATC 240

Qy 246 ProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 265

Db 241 CCCCTGGAGACCAAAACCTTAGGCAAGCAGCCATTCCTAGGGACATTCAGGACAACTC 300

Qy 266 IleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPhe 285

Db 301 ATTGAGATGGACATTTCTCGGCGCTCCCGCCATGATGGGGTTACAGTACGGGCACTTC 360

Qy 286 LeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305

Db 361 CTCTTCAAGCCTGGAGGCACCTCCCGCTCTTCTGCACCAATCACAGGGTACCCACTG 420

Qy 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325

Db 421 ACGTCCAGCACAGTGTACTCTCTCCCGCGCCAGACCCCTGCCCGCGAGCACTTCGCCCGG 480

Qy 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345

Db 481 CCGGCTTTAACTCAAGAGCCCTCCAAGTACTGTAAGTGAAGTGGCAGCCCTGAGC 540

Qy 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365

Db 541 GCCATCGTCATCTCAGCCACTCTGGTCATCTGCTGCTGGCATACTTTGTGGCCATGCACCTG 600

Qy 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAsp 385

Db 601 TTTGGCCTAACTGGCACCTGCAGCCCGATGGAGGGGAGATGTATGAGATCACGGAGGAC 660

Qy 386 ThrAlaSerSerTrpProValProThrAspValSerLeuTyrPro 400

Db 661 ACAGCCAGCAGTTGGCCTGTGCCAACCGACGCTCTCCCTATATACCCC 705

RESULT 7

US-09-623-326-6

; Sequence 6, Application US/09623326

; Patent No. 6607883

; GENERAL INFORMATION:

; APPLICANT: Frey et al.

; TITLE OF INVENTION: Polymerase Chimerae

; FILE REFERENCE: 4894

; CURRENT APPLICATION NUMBER: US/09/623,326

; CURRENT FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: DE 198 10 879.6

; PRIOR FILING DATE: 1998-03-13

; NUMBER OF SEQ ID NOS: 44

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 2949

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: polynucleotide

US-09-623-326-6

Alignment Scores:

Pred. No.:	0.0321	Length:	2949
Score:	131.50	Matches:	101
Percent Similarity:	37.61%	Conservative:	31
Best Local Similarity:	28.77%	Mismatches:	115
Query Match:	6.17%	Indels:	104
DB:	4	Gaps:	20

SEQ14-X-AT-28-64-76 (1-400) x US-09-623-326-6 (1-2949)

QY	3	VallysGluArgLysProTyrArg	-----SerLeuThrArgArg	16
Db	1665	GTCAAGCACAGGAACCTTGTAGAGTGGTTCTTACTTAGAAAGCCTACGAAAGAACGA	1724	
QY	17	AspAlaGluArg	-----ArgTyrThrSerSerSerAla***SerGluGluGly	32
Db	1725	AGTAGCTCCAAACAAGCCAAGTGAAGAGGAGTATCAAAGAAGGCTCAGGGAGAGCTACAC	1784	
QY	33	-----LysAlaProGlnLys-SerTyrSerSerSerGluThrLeuLysAlaTyrAs	49	
Db	1785	AGGTGATTCGTGCGCCTGGACGTGGCCTATCTCAGGGCCTTGTCCCTGGAGGTGGCCGA	1844	
QY	49	pGlnAspAlaArgLeuAlaTyrGlySerArgVallysAspIle***ProGlnGluAlaG1	69	
Db	1845	GGAGATGCCCCGCCTC	-----GAGGCCGA	1868
QY	69	uGluPheCysArgThrGly**AsnPheThrLeuArgGluLeuGly--LeuGluGluVa	88	
Db	1869	GGTCTTCGCCTGGCCGGCCACCCCTTCAACCTCAACTCCCGGGACCAGCTGGAAAGGGT	1928	
QY	88	lThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSe	108	
Db	1929	C-----CTCTTT---GACGAGTAGGGCTTCCCGCCATCGGCAAGAC	1967	
QY	108	rMetGlyAlaGly-----SerAspAlaAspMetGluAlaAspThrValLeuSe	124	
Db	1968	GGAGAAGACCGGCAAGCGCTCCACCAGCGCGCGCTCTTGAGGCC-----CTCCG	2018	
QY	124	rProGluHisProValArgLeu-TrpGlyArgSerThrArgSerGlyArgSerSerCysL	144	
Db	2019	CGAGGCCACACCC-----ATCGTGGAGAAGATCCTGCAGTACCGGGAGCTACCA	2068	
QY	144	eSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGluT	164	
Db	2069	--AGCTGAAGAGCACCT-----ACATTGACCCCTTGCCGGACCTCATCCACC	2113	
QY	164	hrAspHisProGly-GlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeu	183	
Db	2114	-----CCAGGACGGG-----CCGCCTCCA--	2132	
QY	184	SerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGly	203	
Db	2133	---CACCCGCTTCAACGACGCGCCACGGCCACGGCCACGGCGAGGCTAAGTAGCTC	2180	
QY	204	AsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGlu	223	
Db	2181	-----CGATCCCAACCTCCAGAACATCCCCGTCGCGCACCCCGCTTGGGAGAG	2228	
QY	224	ProProAlaGlyGlyAlaGlnGluProAlaHis-AlaGlnGluAsnTrpLeuLeuAsnSe	243	
Db	2229	GATCCGCGGGC-----CTTCATCGCCGAGGAGGGGTGGTATTGGTGGC	2273	
QY	243	rAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAs	263	
Db	2274	C-----CTGGACTA	2282	
QY	263	pAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspG1	283	
Db	2283	TAGCCAGATAGACTCAGGTGCTGGCC-----CACCTCTCCGGCGACGAGAACCT	2333	
QY	283	yHisPheLeuPhe-LysProGlyGlyThrSer-----ProLeuPheCysThrT	299	

Db	2334	GATCCGGTCTTCCAGGAGGGCGGGACATCCACAGGAGACCGCCAGCTGGATGTT	CGG	2393	
QY	299	hrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLeuP	319		
Db	2394	CGTCCCCGGGAGGCCGTGGACCCCTGATGCGCCGGCGGCCAAGACCATCAACTTCGG	2453		
QY	319	roArgSerThrPheAlaArgPro	326		
Db	2454	GGTCTCTACGGCATGTCGGCCC	2476		
RESULT 8					
US-09-231-899-69					
; Sequence 69, Application US/09231899					
; Patent No. 6566583					
; GENERAL INFORMATION:					
; APPLICANT: Lassner, Michael					
; APPLICANT: Metz, James G					
; APPLICANT: Facciotti, Daniel					
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKS GENES					
; FILE REFERENCE: CGNE.131.02US					
; CURRENT APPLICATION NUMBER: US/09/231,899					
; CURRENT FILING DATE: 1999-01-14					
; EARLIER APPLICATION NUMBER: 60/048,650					
; EARLIER FILING DATE: 1997-06-04					
; EARLIER APPLICATION NUMBER: 09/090,793					
; EARLIER FILING DATE: 1998-06-04					
; NUMBER OF SEQ ID NOS: 86					
; SOFTWARE: PatentIn Ver. 2.0					
; SEQ ID NO 69					
; LENGTH: 4446					
; TYPE: DNA					
; ORGANISM: Schizochytrium aggregatum					
US-09-231-899-69					
Alignment Scores:					
Pred. No.:		0.0614	Length:	4446	
Score:		131.50	Matches:	103	
Percent Similarity:		32.72%	Conservative:	39	
Best Local Similarity:		23.73%	Mismatches:	167	
Query Match:		6.17%	Indels:	125	
DB:		4	Gaps:	19	
SEQ14-X-AT-28-64-76 (1-400) x US-09-231-899-69 (1-4446)					
QY	21	ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer	40		
Db	3037	CGCGAGGTGGCAGCTCTATTGCTGCGATTGAAGCGCTCGGGCGCAAGGCCATCTACTCG	3096		
QY	41	SerSerGluThrLeuLysAlaTyrAsp-----GlnAspAlaArgLeuAla	55		
Db	3097	TCGTGCGACGTGAACCTCTGCCGCCGACGTGGCCAAAGCGCGTGGCGATGCCAGTCCACG	3156		
QY	56	TyrGlySerArgValLysAspIle-----	63		
Db	3157	CTCGGTGCGCGCTCTCGGGCATCGTTTCATGCCTCGGGCGTGTCCGCGACCGTCTCATC	3216		
QY	64	-----***ProGlnGluAlaGluGluPheCysArgThrGly**AsnPheThrLeu	80		
Db	3217	GAGAAGAAGCTCCCCGACGAGTTCGACGCCGTC-----TTTGGCACC	3258		
QY	81	ArgGluLeuGlyLeuGluGluVal-ThrProProHisGlyThrLeuTyrArgThrAspI1	100		
Db	3259	AAGGTACCGGTCTCGAGAACCTCTCTGCGCGCGC-----TCGAC	3297		
QY	100	eGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAs	120		
Db	3298	CGCGCCAAACCTCAAGCACATGGTCTCTTCAGCTCGCTCGCCGCTTCCACGGCAACGTC	3357		
QY	120	pThrValLeuSer-----ProGluHisProValArgLeuTrpGly-----	133		
Db	3358	GGCCAGTCTGACTACGCCATGGCCAAACGAGGCCCTTAAACAAGATGGGCTCGAGCTCGCC	3417		

QY	134	-ArgSerThrArgSerGlyArgSerSerCysLeuSer-SerArgAlaAsnSerAsnLeuT	153
Db	3418	AAGGACGTCCTCGTCAAGTCGATCTGCTTCGGTCCCTGGACGGTGGCATGTGTACGCCG	3477
QY	153	hrLeuThrAspThrGluHisGluAsnThrGluThrAspHisPro	167
Db	3478	CAGCTCAAGAAAGCAGTCCAGGAGATGGGCGTGCAGATCATCCCCCGGAGGGCGGCGCT	3537
QY	168	-----GlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProProLeuSerH	185
Db	3538	GATACCGTGGCGG-----CATCGTGTCTCGGCTCCTCGCCGGCTGAGATCCTTGTTCGGC	3591
QY	185	isAlaHisThrProAsnGln-----HisHisAlaAla	195
Db	3592	AACTGGCGCACCCCGTCCAAAGAAGGTGGCTCGGACACCATCACCTGCACCGCAAGATT	3651
QY	196	-----SerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProA	214
Db	3652	TCCGCCAAGTCCAAACCCCTTCCTCGAGGA-----CCACGT---CATCCAGGGCCGC	3699
QY	214	laProThrAspHis-----	218
Db	3700	CGCGTGTGCCCATGACGTGGCCATTGGCTCGCTCGGGAGACCTGCCTCGGCCCTCTTC	3759
QY	218	-----	218
Db	3760	CCGGGCTACTCGCTCTGGGCCATTGACGACGCCCGCAGCTCTTCAAGGTGTCACTGTCTGCAC	3819
QY	219	-----SerLeuSerGlyGluPro---ProAlaGlyGlyAla-----GlnG	231
Db	3820	GGCGACGTCAACTCGGAGGTGACCTTCACCCCGTCGACGGCGCCCTCGGGCGCGGTCAAC	3879
QY	231	luProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArg-	250
Db	3880	GTCCAGGCCACGCTCAAGACCTTTCCAGCGGCAAGCTGGTCCCGCCTACCGCGCGCTC	3939
QY	251	-----AsnLeuGlyLysGlnPropheLeuGlyThrLeuGlnAspAsnLeuIle	266
Db	3940	ATCGTGTCTCCAACACGAGGGCGGCCCGGCCAACGCCACCATGCAGCGCCCTCGCTC	3999
QY	267	GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeu	286
Db	4000	GATGCCGAT-----CCGGCGCTCCAGGGCTCCGCTCTACGACGGCAAGACCCCTC	4047
QY	287	PheLys---ProGly-GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLe	305
Db	4048	TTCCACGGCCCGGCCCTTCGGCGGCATCGATGACGTGCTCTCGTGACCAAGAGCCAGCTT	4107
QY	305	uThrSerSerThrValTyrSerProProProArgPro-----LeuProArgSe	321
Db	4108	GTGGCCAAAGTGCAGCGCTGTCCCGGCTCCGACGCCCGCTCGCGGGCGAGTTTGCCACCGAC	4167
QY	321	rThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCy	341
Db	4168	ACTGACGCCCATGACCCCTTCGTGAACGACCTGGCCCTTTCAGGCCATGCTCGTCTGGGTG	4227
QY	341	sAlaAlaLeuSerAlaIleValIleSerAlaThr	352
Db	4228	CGCGCACGCTCGGCCAGGTCGGCTCGGCTCCCAACT	4261

RESULT 9

US-09-051-994-1/c	
; Sequence 1, Application US/09051994A	
; Patent No. 6602683	
; GENERAL INFORMATION:	
; APPLICANT: REGENERON PHARMACEUTICALS, INC.	
; TITLE OF INVENTION: BIOLOGICALLY ACTIVE EPH FAMILY LIGANDS	
; FILE REFERENCE: REG-341-PCT-US	
; CURRENT APPLICATION NUMBER: US/09/051,994A	
; CURRENT FILING DATE: 1998-04-24	
; EARLIER APPLICATION NUMBER: PCT/US96/17201	
; EARLIER FILING DATE: 1996-10-25	
; EARLIER APPLICATION NUMBER: 60/007,015	

Qy	106	GlyTyrSerMetGlyAlaGlySerAspAlaAspMetGlu---	AlaAspThrValLeuSer	124
			:::	
Db	1351	GGGAGGAGTGAAGGGCTGGGGGGGGCACAGGGGGCGAGATGTTAGAGGTGTCCTCAAA	1292	
Qy	125	Pro-----GluHisProValArgLeuTrp-----	GlyArg	134
			:::	
Db	1291	CCAGGAATTAAACTGGAGGAGCACCCCAAGAGGGTGGATT	CAGGATAGCCACGTGAGA	1232
Qy	135	SerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsn---	LeuThr	153
			:::	
Db	1231	GGAGCCCTCATACCTTGTAGTAGATGTTTGGAGGGCTCTGGGGGGGCCCATCCTGCACGA	1172	


```
QY 63 e***ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeuArgGluLe 83
Db 5829 CTCGCCGGGACGGCGGGCCCTGCAGAGACAGCCCC-----5790

QY 83 uGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIleGlyLeuPr 103
Db 5789 -----GCCGGCGCGCGCGGA-GAGGGCGCCCTCCGGCGGCGCG-----5749

QY 103 oGlnCysGlyTyrSerMetGlyAla-----GlySerAspAlaAspMetGlu---AlaAs 120
Db 5748 -----GGCTTCAGCAGTCCAGCAGCGCGCGCAGCGACTCGGACCTCTCCCGGCGCG 5696

QY 120 pThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyAr 140
Db 5695 CTCGCCCTCCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCGGCCTC 5636

QY 140 gSerSerCysLeuSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGl 160
Db 5635 CTCCTCG-----TCCTCGTCTCTCTCGTCTCTCTCTCTCTCTCTCTCTCGAGGGAGA 5582

QY 160 uAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPr 180
Db 5581 AGAACAGAGAGAGTCCGGCCCGCGCGCG-----CCACTCGCGCGCGCGCGCGCGCG 5528

QY 180 oProProLeuSerHisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLe 200
Db 5527 GTCTCGCGCGCGCGCGCGCGCGCGCGCG-----CCCTCGGCCTCTCTCCGC 5480

QY 200 uAsnArgGlyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLe 220
Db 5479 CTCGCCACCTCTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCG-----5434

QY 220 uSerGlyGluPro--ProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrp 239
Db 5433 -----GAGCCCGCGCGCGCGCGCGCAGGAGCGCGCTCCACCAACACCACTCTC 5381

QY 240 LeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnPropheLeuGly 259
Db 5380 GCTCATGGCCGACGGGCGCGCGCGCGGAC--CGACGGCGCGCTGTCTACCCCGCTCGGGGA 5324

QY 260 ThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAla 279
Db 5323 GCCCTGGCC-----5315

QY 280 TyrSerAspGlyHisPheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThr 299
Db 5314 -----CGGCTCCGACCTCCGGCCGACGGCGCGG-TCCGCTACGGCG 5274

QY 300 SerProGlyTyrProLeuThrSerSerThrValTyrSerProProProArgProLeuPro 319
Db 5273 GCGCGGGGACTCCCGCGAGGGGCTCTGGGACGAGGACGACGTCCGCCAGGCGCGCGGCC 5214

QY 320 ArgSerThrPheAlaArgProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrp 339
Db 5213 ---GCTACCGCGCGCGCGCGCGG-----CCGTGCCCGTCTTCATCCCGGAGATGG 5166

QY 340 -----LysCysAlaAlaLeuSerAla 346
Db 5165 GGGACTCGAGGAAGCAGCAGCAGGCGCCTCGTGGCCTCATCTACAGCG 5118
```

```
RESULT 11
US-09-095-443-1
; Sequence 1, Application US/09095443
; Patent No. 6342593
; GENERAL INFORMATION:
; APPLICANT: Plowman, Gregory
; APPLICANT: Peles, Eior
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF ALP RELATED DISORDERS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
```

```
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,443
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/049,477
; FILING DATE: June 12, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 235/055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4456 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-095-443-1

Alignment Scores:
Pred. No.: 0.146 Length: 4456
Score: 127.50 Matches: 93
Percent Similarity: 32.99% Conservative: 37
Best Local Similarity: 23.60% Mismatches: 162
Query Match: 5.98% Indels: 103
DB: 4 Gaps: 16

SEQ14-X-AT-28-64-76 (1-400) x US-09-095-443-1 (1-4456)

QY 2 AspValLysGluArgLysProTyrArgSer-----LeuThrArgArg 15
Db 1357 GAGCTGAAGAAGAAGCGCGCGCCACAGCCCCAAAGCGGTGCTGCCCGCAGG 1416

QY 16 ArgAspAlaGluArgArgTyrThrSerSerAla**SerGluGluGlyLysAlaPro 35
Db 1417 GAGGAGAGTGAG-----GCAGTGAAGCAGGAGACCCCCCT 1452

QY 36 GlnLysSerTyrSerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAla 55
Db 1453 -----GAGGAGCTGGCAGCCTCCCCCCTGAC-----ATGGTG 1485

QY 56 TyrGlySerArgValLysAspIle***Pro-----65
Db 1486 GCTGCCACGACTGCTGTGACAC-CTTCTGGGAAGTGCCACCCCGCTCCACTTTCCTCC 1544

QY 66 GlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeuArgGluLeuGlyLeu 85
Db 1545 CAGCCCCCTTCCCAGCTCCACAGGCCCGCAGGACCCCACTATCTCTCAGGCCCCCTTGGCCCC 1604

QY 86 GluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCys 105
Db 1605 TGGTACCTACTCGGCGCCAC-----1625

QY 106 GlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeuSerPro 125
Db 1626 -----CCAGCTGATACAGCCCGGCCCC 1649

QY 126 GluHisProValArgLeuTrpGlyArgSerThrArgSerGly-----ArgSerSerCys 143
```



```
Db 1650 AGGGCCCCATGCAAT-----GCCCGTAGCACCTGGGCCTGCCCTCTACCCAGCCCCCTGC 1703
QY 144 LeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGlu 163
Db 1704 CTACACACGGAGCTGGCCTTGTCGCCCGATCCTCCCAACAGCATGGCGTGGT--GAG 1760
QY 164 ThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeu 183
Db 1761 CAGTCCCTATGTGGGGTAGGGCCGCCACACAGTTGCAGGTCTCCCTCGGCCCCACC 1820
QY 184 Ser-----HisAlaHisThrProAsnGlnHisHisAlaAla 195
Db 1821 TCCTCAATTCTCAGGCCCCGAGTTGGCCATGGCGGTTCGGCCAGCCACCACAGTAGA 1880
QY 196 -SerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAlaPr 215
Db 1881 TAGCATCCAGCGGCCCATCCCCAGCCACAGCCCCACGGCCAAACCCACCCCTGCTCC 1940
QY 215 oThrAspHisSerLeuSerGlyGluProPro-----Al 226
Db 1941 TCCCGCGCCTGCTTCCCTGTGCCCCACCGCAGCCACTGCCACGCCTTACACCTACCC 2000
QY 226 aGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrp-LeuLeuAsnSerAsnIleP 246
Db 2001 TGCAGGGGCTAAGCAACCCATCCACAGCACAGCACCACTTCTCTCTGGGATCCCCACAG 2060
QY 246 roLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeu 266
Db 2061 TTTTCCAGCCCCAAGGATTGGGCCCCAGCCCC-----AGCCCCATCCTCAGCCCCCATCC 2114
QY 266 leGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyr---SerAspGlyHisP 285
Db 2115 TTCACAAGCGT--TTGGGCTCAGCCCCACAGCAGCCCCCTTCCACTCCAGCATCCACA 2171
QY 285 he-LeuPheLysProGlyGlyThrSerProLeu-----PheCys--- 297
Db 2172 TCTCTTCCACCCCGAGCCCCCAGGACTCCTACCCCCACAATCCCCCTACCCCTATGCCCC 2231
QY 298 -----ThrThrSerProGlyTyrProLeuThrSerSerThr 309
Db 2232 TCAGCCTGGGTCCTGGGGCAGCGGCCACCCCTACACACCCAGCTCTACCCAGGTCC 2291
QY 310 ValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsn 329
Db 2292 CGCTCAAGACCCTCTGCCAGCCCACTCAGGGGCTCTGCCTTCCCCAGCCCTGGGCCCCC 2351
QY 330 LeuLysLysProSerLysTyrCysAsnTrpLysCys 341
Db 2352 TCAGCCTCCCCATCCCCCACTGGCATATGGTCCTGC 2387
```

```
RESULT 12
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
```

```
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 9.92e+03 Length: 4403765
Score: 126.50 Matches: 79
Percent Similarity: 36.42% Conservative: 43
Best Local Similarity: 23.58% Mismatches: 121
Query Match: 5.93% Indels: 92
DB: 3 Gaps: 16

SEQ14-X-AT-28-64-76 (1-400) x US-09-103-840A-2 (1-4403765)

QY 32 GlyLysAlaProGlnLysSerTyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp 51
Db 1334300 GGTCCGGTCCCAATCTGCATCAGGGTGCTCAAATCTCTTCAC----- 1334256
QY 52 AlaArgLeuAlaTyrGlySerArgValLysAspIle**ProGlnGluAlaGluGluPhe 71
Db 1334255 -----CACGGCCAAAGCCACGAGAT-----GACTTT 1334229
QY 72 CysArgThrGly**AsnPheThrLeuArgGluLeuGlyLeuGluValThrProPro 91
Db 1334228 CGCCGGCAAGTGCAACAACCTCGCGTCGGGAAGCCGGGATACGACGTGT-----CCC 1334175
QY 92 HisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAla 111
Db 1334174 CGTGGG-----AGAACGGGA-----CGATGTGGTCTCGGTCTCCGTGCC 1334136
QY 112 GlySer-AspAlaAspMetGluAlaAspThrValLeuSerProGluHisProVal-ArgL 131
Db 1334135 ACCACGGACGGGACTTTTCAC-----CTCGTCCAGCCGAATCCCACTCGCGGG 1334085
QY 131 euTrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSer-----SerArgA 148
Db 1334084 CAAAGCGGATGACATCGGCGGAACGGCGCAGCGAGCTGTTCGACTACCGTTGAGCAGAT 1334025
QY 148 laAsnSerAsnLeuThrLeu-ThrAspThrGluHisGluAsnThr----- 162
Db 1334024 CGTCGAGGAACATCGCCTTGAACCTCGGCGCAGCCAGCAAAATGCCGTGCGCCGGGTG 1333965
QY 163 -----GluThrAspHisProGlyGlyLeu 170
Db 1333964 AGACAGGCCCATACAGGTCGAGGGCAGGGGACGCGACGGGCGGGCCCGGATCAGCA 1333905
QY 171 GlnAsnHisAlaArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsn 190
Db 1333904 AGCTCGCACCCAGCGCGAGGGGGTGCAGCCCACTGACGAGCGGCGCCACCGCCGAAC 1333845
QY 191 Gln-HisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAs 210
Db 1333844 CAAGCGCATCAAAACCGCCGCTAATCGCTCCGGGCCCGCGGT-----CGGCGCGACGC 1333791
QY 210 nProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaG 230
Db 1333790 CGCCGAGGACACCGGGCGCGACCAACCCGGTC---GGGACGCC-----GGCGCGCAGC 1333740
QY 230 nGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrAr 250
Db 1333739 CCAGGGTGTATGGGCCCGCCCGCCGACAGGCCCAACACGGCCCATCTT---GTCGATGCCGA 1333683
QY 250 gAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspI 270
Db 1333682 GCGTGTCCGCGATGTGTCG-----CAGAT 1333659
QY 270 eLeuGlyAlaSerArgHisAspGly-AlaTyrSerAspGlyHisPheLeuPheLysPro 290
Db 1333658 CGTCGGC---GAACGCCAAGATGGTTTCGTACTGATCGGCGCTCGAGGCGCCGATGCCGG 1333602
QY 290 lyGlyThrSerProLeu----- 295
Db 1333601 GCCGATCGACGCCCAATCAGACGAATATTGTGTGTCGGCGTAGACCCGGGCTCGGTGCG 1333542
```

QY 296 --PheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProp 315
Db 1333541 GGATCTGCCGCGCGGCTTGGGTGCCATGCAGCCAGAACCGCA----- 1333495

QY 315 roArgProLeuProArgSerThrPheAlaArgPro 326
Db 1333494 --CGACCCTGCGGGCGCGCAACTCGGCGAAGCCG 1333462

RESULT 13
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 9.95e+03 Length: 4411529
Score: 126.50 Matches: 79
Percent Similarity: 36.42% Conservative: 43
Best Local Similarity: 23.58% Mismatches: 121
Query Match: 5.93% Indels: 92
DB: 3 Gaps: 16

SEQ14-X-AT-28-64-76 (1-400) x US-09-103-840A-1 (1-4411529)

QY 32 GlyLysAlaProGlnLysSerTyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp 51
Db 1334832 GGTCGGGTCCCAATCTGCATCAGGGTGCTCAAAATCTCTTAC----- 1334788

QY 52 AlaArgLeuAlaTyrGlySerArgValLysAspIle**ProGlnGluAlaGluGluPhe 71
Db 1334787 -----CACGGCCAAAGCCAGCGAGAT-----GACTTT 1334761

QY 72 CysArgThrGly***AsnPheThrLeuArgGluLeuGlyLeuGluValThrProPro 91
Db 1334760 CGCCGGGCAAGTGCACAACTTCGCTCGGAAGCCGGGATACGAGTGT-----CCC 1334707

QY 92 HisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAla 111
Db 1334706 CGTGGG-----AGAACGGGA-----CGATGTGGTCTCGGTCTCCGTGCC 1334668

QY 112 GlySer-AspAlaAspMetGluAlaAspThrValLeuSerProGluHisProVal-ArgL 131
Db 1334667 ACCAGCGGACGGGACTTTTAC-----CTCGTCCAGCCGGAATCCCGAGTCGCGG 1334617

QY 131 euTrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSer-----SerArgA 148
Db 1334616 CAAAGCGGATGACATCGGCAACGGCGCAGCGAGCTGCTTGGGACTACCGTTGAGCAGAT 1334557

QY 148 laAsnSerAsnLeuThrLeu-ThrAspThrGluHisGluAsnThr----- 162
Db 1334556 CGTCGAGGAACATCGCCTTGAACCTCGGGCCGAGCCAGCAATGCCGGTCGGCCCGCGGTG 1334497

QY 163 -----GluThrAspHisProGlyGlyLeu 170
Db 1334496 AGAGCAGGCCATACAGGTTCGAGGGCAGGGGACGGGCGCGCGCGGATCAGCA 1334437

QY 171 GlnAsnHisAlaAargLeuArgThrProProProProLeuSerHisAlaHisThrProAsn 190
Db 1334436 AGCTGCACCCAGCCGACGGGGTGTCCGCCACCTGCAGAGCGGCGCCACCGCGCAAC 1334377

QY 191 Gln-HisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAs 210
Db 1334376 CAAGGCGCATCAAAACCGCGCTAATCGCTCCGGGCGCGCGGT-----CGGCGCGACGC 1334323

QY 210 nProSerProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGl 230
Db 1334322 CGCCGAGGACACCGGGCGGCGCACCCCGT-----GGGCGCGCACG 1334272

QY 230 nGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrAr 250
Db 1334271 CCAGGGTGTATGGCCCCCGCCGACAGGCCACCCATCTT---GTCGATGCCGA 1334215

QY 250 gAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAsnLeuIleGluMetAspIl 270
Db 1334214 GCGTGTCCGGATGGTCCG-----CAGAT 1334191

QY 270 eLeuGlyAlaSerArgHisAspGly-AlaTyrSerAspGlyHisPheLeuPheLysProG 290
Db 1334190 CGTCGGC---GAACGCCAAGATGGTTCTGACTGATCGGCGCTCGAGGCGCGGATGCCGG 1334134

QY 290 lyGlyThrSerProLeu----- 295
Db 1334133 GCCATCGACGCCAATCAGACGAATATTGGTGTCTCGGCGTAGACCCGGGCTTCGGTGC 1334074

QY 296 --PheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProp 315
Db 1334073 GGATCTGCCGCGCGGCGCTTGGGTGCCATGCAGCCAGAACCGCA----- 1334027

QY 315 roArgProLeuProArgSerThrPheAlaArgPro 326
Db 1334026 --CGACCCTGCGGGCGCGCAACTCGGCGAAGCCG 1333994

RESULT 14
US-08-471-066B-1
; Sequence 1, Application US/08471066B
; Patent No. 5837450
; GENERAL INFORMATION:
; APPLICANT: Dahlberg, James E.
; APPLICANT: Lyamichiev, Victor I.
; APPLICANT: Brow, Mary Ann D.
; TITLE OF INVENTION: 5' Nucleases Derived From Thermostable
; TITLE OF INVENTION: DNA Polymerase
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,066B
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/254,359
; FILING DATE: 06-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/073,384
; FILING DATE: 04-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/986,330
; FILING DATE: 07-DEC-1992

```

;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ingolia, Diane E.
; REGISTRATION NUMBER: 40,027
; REFERENCE/DOCKET NUMBER: FORS-01800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-471-066B-1

Alignment Scores:
Pred. No.: 0.0808 Length: 2507
Score: 126.00 Matches: 98
Percent Similarity: 37.13% Conservative: 29
Best Local Similarity: 28.65% Mismatches: 123
Query Match: 5.91% Indels: 92
DB: 2 Gaps: 19

SEQ14-X-AT-28-64-76 (1-400) x US-08-471-066B-1 (1-2507)

QY 10 ArgSerLeuThrArgArgArgAspAlaGluArgTyrThrSerSerSerAla***Ser 29
Db 1273 AGGCTCCITGGCTTTACCGGAGGTGAGAGGCCCTTTCCGCTGCTCCTGGCCACATG 1332
QY 30 Glu---GluGlyLysAlaProGlnLysSerTyrSerSerGluThrLeuLysAlaTyr 48
Db 1333 GAGGCCACGGGGTGGCCTGGACGTGGCCTATCTCAGGGCCTTGTCCCTGGAGGTGGCC 1392
QY 49 AspGlnAspAlaArgLeuAlaTyrGlySerArgVallysAspIle***ProGlnGluAla 68
Db 1393 GAGGAGATCGCCCGCCTC-----GAGGCC 1416
QY 69 GluGluPheCysArgThrGly***AsnPheThrLeuArgGluLeuGly---LeuGluGlu 87
Db 1417 GAGGTCTTCGCCTGGCCGCGCCACCCCTTCAACCTCAACTCCCGGGACCAGCTGGAAGG 1476
QY 88 ValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeuProGlnCysGlyTyr 107
Db 1477 GTC-----CTCTTT---GACGAGCTAGGCTTCCGCCCATCGGCAAG 1515
QY 108 SerMetGlyAlaGly-----SerAspAlaAspMetGluAlaAspThrValLeu 123
Db 1516 ACGGAGAAGACCGGCAAGCGCTCCACCAGCGCGCTCTCTGGAGGCC-----CTC 1566
QY 124 SerProGluHisProValArgLeu-TripGlyArgSerThrArgSerGlyArgSerSerCy 143
Db 1567 CGCGAGGCCACCCC-----ATCGTGAGAGAGATCCTGCAGTACCGGAGCTCACCA-- 1618
QY 143 sLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGl 163
Db 1619 ----AGCTGAAGAGCACCT-----ACATTGACCCCTTGCCGGACCTCATCCACC-- 1663
QY 163 uThrAspHisProGly-GlyLeuGlnAsnHisAlaArgLeuArgThrProProProL 183
Db 1664 -----CCAGGACGGG-----CCGCTCCA- 1682
QY 183 euSerHisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgG 203
Db 1683 -----CACCCGCTTCAACAGACGCGCCACGGCCACGGCAGGCTAAGTAGCTC----- 1730
QY 203 lyAsnPheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyG 223
Db 1731 -----CGATCCCAACCTCCAGAACATCCCGCTCCGCACCCCGCTGGGCAG 1776
QY 223 luProProAlaGlyGlyAlaGlnGluProAlaHis-AlaGlnGluAsnTrpLeuLeuAsn 242
Db 1777 AGGATCCGCGGGC-----CTTCATCGCCGAGGAGGGGTGGCTATTGGTG 1821
```

```

QY 243 SerAsnIleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGln 262
Db 1822 GCC-----CTGGAC 1830
QY 263 AspAsnLeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAsp 282
Db 1831 TATAGCCAGATAGAGCTCAGGGTGTGGCC-----CACCTCTCCGGCGACGAGAAC 1881
QY 283 GlyHisPheLeuPhe-LysProGlyGlyThrSer-----ProLeuPheCysTh 298
Db 1882 CTGATCCGGGTCTTCCAGGAGGGCGGACATCCACACGGAGACCGCCAGCTGGATGTTT 1941
QY 298 rThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProProArgProLe 318
Db 1942 GGCCTCCCGGAGGCCGTGGACCCCTGATGCGCGCGGCCCAAGACCATCACTTC 2001
QY 318 uProArgSerThrPheAlaArgProAlaPheAsnLeuLysLys---ProSerLysTyr 336
Db 2002 GGGGTCTCTACGGCATGTGCGGCCACCGCCTCTCCCAGGAGTAGCATCCCCCTTAC 2059
RESULT 15
US-09-252-991A-11748
; Sequence 11748, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11748
; LENGTH: 1998
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11748
```

```

Alignment Scores:
Pred. No.: 0.07 Length: 1998
Score: 125.00 Matches: 85
Percent Similarity: 33.42% Conservative: 40
Best Local Similarity: 22.73% Mismatches: 128
Query Match: 5.86% Indels: 122
DB: 4 Gaps: 17
```

SEQ14-X-AT-28-64-76 (1-400) x US-09-252-991A-11748 (1-1998)

```

QY 5 GluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArgArgTyrThrSer 24
Db 716 CAGCGCCAGACGCTGCAAGGCCTGGAGCGCGCAACGACGTCGCTGCGGACAGTTGACC 775
QY 25 SerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSerSerSerGluThr 44
Db 776 GAGCGCCGCAACGAACTCGCCGGGCTTTCCGCGCGCCAGGCCAACAGCTCGCGGAAACC 835
QY 45 -----LeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVallys 61
Db 836 CGCCGCCAGCTCAGCGCGGTGGAGCAGGAC-----CTGGCCGAAAGCAAGCAAGCGC 889
QY 62 AspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeuArg 81
Db 890 ACCTTGCTGTCAACCGCGCGGAGAGCGGATCGCCACCGCC-----GTGCTCGCC 940
QY 82 GluLeuGly-----LeuGluGlu-----LeuGluGluValThr 89
Db 941 GAAGCCGGGACAGCCGTCGACAGCTCGCGTCCGCTGCTGAGCATCGTTCGCCCGACACC 1000
```


QY 90 ProProHisGlyThrLeuTyr-----ArgThrAspIleGlyLeuProGlnCysGlyTyr 107
|||
Db 1001 CCGTTGCAGGCGGAAGTCTACGGCCGAGCAAGTCCATCGTTTCATCCGGCCGGGC--- 1057
|||
QY 108 SerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrValLeu-SerProGluHi 127
|||
Db 1058 -----GACGGGTGCTGATCCGCTACCAG 1081
|||
QY 127 sProValArgLeuTrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerAr 147
|||
Db 1082 GCCTATCCGTACC-----AGAAGTTCGGCCAGTACCACGGCAAGTGCAGTCGATCTCC 1135
|||
QY 147 gAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsnThrGluThrAsp----- 165
|||
Db 1136 CCGGC--CAGCGTCTCCTATGCCGAGCTTTCAGCATGTGTCGGCGGCGTACCGGGGCTCG 1193
|||
QY 166 -----HisProGlyGly----- 169
|||
Db 1194 GCCAGGATGGCGAGCAGCTGTACCGCTGCGGGTAACCTTCGACGACCAGCGGTGACCG 1253
|||
QY 170 -----LeuGlnAsnHisAlaArgLeuArgThrProPro-- 181
|||
Db 1254 CTTACGGCCAGCGCGTCCGCTGCAGAGCGGCATGCTGTGACGCCGACATCCTCCAGG 1313
|||
QY 182 -----ProLeu-----SerHisAlaHisThrProAsnGlnHisAl 194
|||
Db 1314 ACACCGCGCGCTCTACGAATGGTGTGTGAACCGCTCTACAGCTGACCGGCAAACTCT 1373
|||
QY 194 aAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSerProAl 214
|||
Db 1374 AGGAACGACCCATGGCCTTCTCGA-----CGCTCTGCCCTGCGCCTGG 1418
|||
QY 214 aProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluProAlaHi 234
|||
Db 1419 GCCG-----CGCCTGCCGTGGTGTGTGCAGACCGAAGCCA 1454
|||
QY 234 sAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeuGlyLy 254
|||
Db 1455 CCGA-----ATGCGGCCTGSCCTGCCTGGCGATGATGCGCGGCTACCACG 1499
|||
QY 254 sGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGlyAlaSe 274
|||
Db 1500 GCCACCA----- 1506
|||
QY 274 rArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysProGlyGlyThrSerPr 294
|||
Db 1507 -----TACCGGCTGATGGAAGTGGCC 1529
|||
QY 294 oLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSerProPr 314
|||
Db 1530 GCGGCTTCTCCGTATCGCTCAAGGGCATCTCCCTCAAGCAACTGATCCAGACGCCCCACC 1589
|||
QY 314 oProArgPro---LeuProArgSerThrPheAlaArgPro 326
|||
Db 1590 GCCTCGGCCCTGGGTACCCGCGCGGTGAA---GCTCGACCT 1626
|||

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 18:05:13 ; Search time 3336 Seconds
(without alignments)
3580.597 Million cell updates/sec

Title: SEQ14-X-AT-28-64-76
Perfect score: 2132
Sequence: 1 MDVKERKPYRSLTRRRDAER.....EITEDTASSWVPVPTDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFPO_spool/MITRA020/runat_06082004_114101_218/app_query.fasta_1.583
-DB=EST -QFMT=fastop -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=MITRA020 @CGN_1_1_3437 @runat_06082004_114101_218 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	924	43.3	620	14	CB532245	CB532245 756697 MA
2	924	43.3	620	14	CB532612	CB532612 757081 MA
3	884	41.5	566	13	BU072782	BU072782 im47g02.y
4	769	36.1	625	13	BU200362	BU200362 604158648
5	725	34.0	878	13	BU474934	BU474934 603364850
6	720.5	33.8	3447	11	AK047423	AK047423 Mus muscu
7	699	32.8	884	13	BQ735622	BQ735622 AGENCOURT
8	685	32.1	402	14	CB544750	CB544750 AMGNNUC:T
9	680	31.9	789	13	BU474512	BU474512 603760992
10	675.5	31.7	887	13	BU119163	BU119163 603143046
11	657	30.8	4556	11	AK034286	AK034286 Mus muscu
C 12	583.5	27.4	490	9	AL922332	AL922332 AL922332
13	568	26.6	870	10	BG036207	BG036207 602326960
14	519.5	24.4	511	29	AY405291	AY405291 Homo sapi
15	519.5	24.4	511	29	AY405292	AY405292 Pan trogl
16	517.5	24.3	511	29	AY405293	AY405293 Mus muscu
17	512.5	24.0	534	10	BF944080	BF944080 QV3-NN020
18	509	23.9	2716	11	AK031198	AK031198 Mus muscu
C 19	490.5	23.0	880	13	BX760712	BX760712 BX760712
C 20	477.5	22.4	568	10	BF953105	BF953105 QV3-NN020
21	470	22.0	2627	11	AK050784	AK050784 Mus muscu
22	456.5	21.4	643	10	BB657984	BB657984 BB657984
23	448	21.0	717	10	BB654584	BB654584 BB654584
C 24	438.5	20.6	593	29	CE756467	CE756467 tigr-gss-
C 25	433	20.3	284	10	BF388169	BF388169 UI-R-CA1-
26	412	19.3	259	28	CC178145	CC178145 XC549 Bay
27	397	18.6	825	13	BU253342	BU253342 603747714
28	388.5	18.2	843	14	CF225019	CF225019 AGENCOURT
29	369	17.3	864	13	BU172491	BU172491 AGENCOURT
30	362	17.0	397	13	BY010856	BY010856 BY010856
31	362	17.0	704	13	BY733523	BY733523 BY733523
32	355	16.7	643	13	BY723994	BY723994 BY723994
33	331.5	15.5	608	9	AL955817	AL955817 AL955817
34	330.5	15.5	707	13	BU229930	BU229930 603947544
35	329.5	15.5	801	13	BU451304	BU451304 603216940
36	287	13.5	461	13	BU473055	BU473055 603761665
37	282.5	13.3	751	13	BU461850	BU461850 603773816
C 38	273.5	12.8	1060	29	CNS04V97	AL308788 Tetraodon
39	272.5	12.8	745	29	CE434866	CE434866 tigr-gss-
40	269.5	12.6	404	29	CG514853	CG514853 OST68836
C 41	257.5	12.1	386	9	AA702693	AA702693 z190g04.s
42	256.5	12.0	674	28	BZ111628	BZ111628 CH230-230
C 43	248	11.6	474	9	AI753786	AI753786 cr16e09.x
44	245	11.5	1013	12	BM017187	BM017187 603643790
C 45	237	11.1	1125	28	CC310899	CC310899 TAM32-30J

ALIGNMENTS

RESULT 1
CB532245/c
LOCUS CB532245 620 bp mRNA linear EST 16-MAY-2003
DEFINITION 756697 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CB532245
VERSION CB532245.1 GI:29395750
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 620)

AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G., Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross match v0.990329.
Plate: LAM8014 row: K column: 19
Seq primer: TAGAAGGCACAGTCGAGG.
Location/Qualifiers
1. .620
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."
ORIGIN
Alignment Scores:
Pred. No.: 1.11e-57 Length: 620
Score: 924.00 Matches: 172
Percent Similarity: 94.71% Conservative: 7
Best Local Similarity: 91.01% Mismatches: 10
Query Match: 43.34% Indels: 0
DB: 14 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x CB532245 (1-620)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
DB 572 ATGGACGTGAAGGAGAGAGCAATACCGCTCGCTGACCCGCGCGCGAGCGC 513
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 512 CGCTACACCACTCGTCGGCCGACAGCGAAGAGGGCAAGCGCCGCGAGAAGTCCTACAGC 453
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 452 TCCAGCGAGACCCCTCAAGGCCTACGACCAAGGACGCGCGCTCACTTACGGCAGCCGCGTC 393
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
DB 392 AAGGACCTGGTACCAACAGAGGCGGAGAGTTCTGCCGCGCAGGCAACCACTTCAGCCTG 333
QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
DB 332 CGCGAGCTGGGGCTGGGGAGGTGACACCCCGCACGGACGCTGTACCGCACGGACATC 273
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
DB 272 GGCCTGCCCACTGCGGCTACTCCCTGGGGCCAGCTCTGAGGCGCGAGCTGGAGCGGAC 213
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
DB 212 GCCGCACTGTCCCCGAGACCCCGTGGGCTGTGGGGCCGAGCACCGGTCAGGACGC 153
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
DB 152 AGCTCCTGCTGTCCAGCCGGGCCAACTCCAACCTCACGCTCACTGACACGGAGCACGAG 93
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180

DB 92 AACACGGAGACCGATCACCCGGGGGGCTGCAGAACCACTTCTCGACTCCGGACACCGCGC 33
QY 181 ProProLeuSerHisAlaHisThrPro 189
DB 32 CCTCCGCTCTCGCATGCCACACACCCCC 6
RESULT 2
CB532612 620 bp mRNA linear EST 16-MAY-2003
LOCUS 757081 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION CB532612
ACCESSION CB532612
VERSION CB532612.1 GI:29396493
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.
1 (bases 1 to 620)
Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G., Wray,J.E. and Keele,J.W.
A second set of bovine ESTs from pooled-tissue normalized libraries
Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross match v0.990329.
Plate: LAM8014 row: K column: 19
Seq primer: GTAATACGACTCACTATAGG.
Location/Qualifiers
1. .620
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."
ORIGIN
Alignment Scores:
Pred. No.: 1.11e-57 Length: 620
Score: 924.00 Matches: 172
Percent Similarity: 94.71% Conservative: 7
Best Local Similarity: 91.01% Mismatches: 10
Query Match: 43.34% Indels: 0
DB: 14 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x CB532612 (1-620)
QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgArgAspAlaGluArg 20
DB 49 ATGGACGTGAAGGAGAGAGCAATACCGCTCGCTGACCCGCGCGCGAGCGC 108
QY 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
DB 109 CGCTACACCACTCGTCGGCCGACAGCGAAGAGGGCAAGGCCCGCAGAGTCTCTACAGC 168
QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
DB 169 TCCAGCGAGACCCCTCAAGGCCTACGACCAAGGACGCGCGCTCACCTACGGCAGCCGCGTC 228
QY 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
DB 229 AAGGACCTGGTACCAACAGAGGCGGAGAGTTCTGCCGCGCAGGCAACCACTTCAGCCTG 288

QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
|||||
Db 289 CGCAGCTGGGCTGGGGAGGTGACACCCCGCACGGGACGCTGTACCGCAGGACATC 348
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
|||||
Db 349 GGCCTGCCCACTGCGGCTACTCCCTGGCGGCCAGCTCTGAGCGGAGCTGGAGCGGAC 408
QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
|||||
Db 409 GCGGCACTGTCCCCGAGCAGCACCCCGTGGGCTGTGGGGCCGAGCAGCGGTTCAGGACGC 468
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
|||||
Db 469 AGCTCCTGCCTGTCCAGCCGGGCAACTCCAACTCAGCTCACTGACACGGAGCACGAG 528
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
|||||
Db 529 AACACGGAGACCGATCACCCGGCGGCGCTGCAGAACCAATCTCGACTCCGACACCGCG 588
QY 181 ProProLeuSerHisAlaHisThrPro 189
|||||
Db 589 CTTCCGCTCTCGCATGCCCCACACCCCC 615

RESULT 3

BU072782 566 bp mRNA linear EST 27-AUG-2002
LOCUS im47g02.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6038211 5'
DEFINITION similar to TR:014667 014667 PRO-NEUREGULIN-1, GAMMA ISOFORM
PRECUSOR ; mRNA sequence.

ACCESSION

VERSION BU072782.1 GI:22513971
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 566)

REFERENCE

AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)

JOURNAL

COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 339.

FEATURES

source

Location/Qualifiers

1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6038211"
/tissue type="Purified pancreatic islet"
/lab_host="DH10B"
/clone lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'

XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

Alignment Scores:

Pred. No.: 8.16e-55 Length: 566
Score: 884.00 Matches: 162
Percent Similarity: 95.86% Conservative: 0
Best Local Similarity: 95.86% Mismatches: 7
Query Match: 41.46% Indels: 0
DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BU072782 (1-566)

QY 165 AspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSer 184
|||||

Db 57 GATCATCCGGCGGCTGCAGAACCCACGCGGCTCCGACGCGCGCGCTCTCG 116
|||||

QY 185 HisAlaHisThrProAsnGlnHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsn 204
|||||

Db 117 CACGCCACACCCCCAACACAGCACCCACGCGGCTCCATTAACTCCCTGAACCGGGGCAAC 176
|||||

QY 205 PheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluPro 224
|||||

Db 177 TTCACGCCGAGGAGCAACCCAGCCCGGCCCCCGACGACCACTCGCTCTCCGAGAGCCC 236
|||||

QY 225 ProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsn 244
|||||

Db 237 CTTGCCGCGCGGCCAGGAGCCTGCCACGCCCGCCAGGAGAACTGGCTGCTCAACAGCAAC 296
|||||

QY 245 IleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsn 264
|||||

Db 297 ATCCCCCTGGAGACCCAGAAACCTAGGCAAGCAGCATTCCTAGGACATTGCGAGCAAC 356
|||||

QY 265 LeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHis 284
|||||

Db 357 CTCATTGAGATGGACATTCTCGCGCCTCCCGCCATGATGGGGCTTACAGTGACGGCAC 416
|||||

QY 285 PheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrPro 304
|||||

Db 417 TTCTCTTCAAGCCTGGAGGCACCTCCCGCTCTCTGACCCACATCACCAGGTACCCA 476
|||||

QY 305 LeuThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAla 324
|||||

Db 477 CTGACGTCCAGCACAGTGACTCTCCCTCCCGCCGACCCCTGGCCCGCAGCAACTTTCGC 536
|||||

QY 325 ArgProAlaPheAsnLeuLysLysPro 333
|||||

Db 537 CGGGCGGNCCTTAACTCAAGAGCCT 563
|||||

RESULT 4

BU200362

LOCUS

DEFINITION

BU200362 604158648F1 CSEQCHN03 Gallus gallus cdna clone ChEST1010p11 5',
mRNA sequence.

ACCESSION

VERSION BU200362.1 GI:25363728

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 625)

REFERENCE

AUTHORS

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534

BU200362 625 bp mRNA linear EST 25-NOV-2002
604158648F1 CSEQCHN03 Gallus gallus cdna clone ChEST1010p11 5',
mRNA sequence.

BU200362.1 GI:25363728

EST.

Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 625)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,

Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source 1. .625
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST1010p11"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_lib="CSEQCHN03"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site_1: EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 2.42e-46 Length: 625
Score: 769.00 Matches: 149
Percent Similarity: 89.94% Conservative: 12
Best Local Similarity: 83.24% Mismatches: 16
Query Match: 36.07% Indels: 2
DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BU200362 (1-625)

Qy 1 MetaSpVallLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 93 ATGGATGTAAGAAGAAAGGAAACCGTATCGATCTCTGACTCGGCGCGGACACGGAGCGC 152

Qy 21 ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 153 CGGTACACCACTCTTCAGCCGAGAGTGAAGACAGCAAGGCTCCTCAGAAGTCTATAGC 212

Qy 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 213 TCCAGTGAGACCCCTGAAGGCTTATGATCAAGACTCCAGGTTGACCTACAGCAATCGGGTC 272

Qy 61 LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu 80
Db 273 AAAGACATGTGTGCACCGAGGCTGATGAATTCTGCCGAGCAGGAGCAACTTCTCTTTG 332

Qy 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
Db 333 CGGGAGCTGGGTCTTGAAGATGTGACTCCACCCATGGGACTTTGTACCGGACTGATATT 392

Qy 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 393 GGGCTGCCTCACTGTGGCTACTCCATCAGCGTGGCTCAGATGCCGACACGGAGGGAT 452

Qy 121 ThrValLeuSerProGluHis-ProValArgLeuTyrGlyArgSerThrArgSerGlyAr 140
Db 453 GTGGTCATGTACCTGAGCATCCCTGTGAGGCTCTGGGGACGCAACACCAATCCGGACG 512

QY 140 gSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1 160
Db 513 CAGCTCCTGCTTGTGAGCGGGCCAACTCCAACTCACCCTCACCACACGGAGCACGA 572

QY 160 uAsn-ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg 177
Db 573 GAATTACCGAACTGATCATCTCTCCAAATCTTCAAAATCATTTCAAGACTCCGA 625

RESULT 5
BU474934 878 bp mRNA linear EST 30-NOV-2002
LOCUS 603364850F1 CSEQRBN21 Gallus gallus cDNA clone ChEST262c14 5', mRNA
DEFINITION sequence.
ACCESSION BU474934
VERSION BU474934.1 GI:25968511
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 878)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.,
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source Location/Qualifiers
1. .878
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="ChEST262c14"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBN21"
/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 6.51e-43 Length: 878
Score: 725.00 Matches: 139
Percent Similarity: 90.91% Conservative: 11
Best Local Similarity: 84.24% Mismatches: 14
Query Match: 34.01% Indels: 1
DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BU474934 (1-878)

QY	1	MetAspValIysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	257	ATGGATGTAAAGAAAGGAAACCGTATCGATCTCTGACTCGGCGCGACACGGAGCGC	316
QY	21	ArgTyrThrSerSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	317	CGCTACACAGCTCTTCAGCCGAGAGTGAGGACAGCAAGGCTCCTCAGAAGTCCTATAGC	376
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	377	TCCAGTGAGACCTGAAGGCTTATGATCAAGACTCCAGGTTCACCTACAGCAATCGGGTC	436
QY	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80
Db	437	AAAGACATGCTGCACCAGGAGGCTGATGAATTCTGCCGAGCAGGAGCCAACTTCTCTTTG	496
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	497	CGGGAGCTGGGTCTTGAAGATGTGACTCCACCCATGGGACTTGTACCGGACTGATATT	556
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	557	GGGTGCTCACTGTGGCTACTCCATCAGCGCTGGCTCAGATGCCGACACGGAGCGGAT	616
QY	121	ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg	140
Db	617	GTGGTCATGTCACCTGAGCATCCTGTGAGGCTCTGGGGACGCAACACCAATCCGGACGC	676
QY	141	SerSerCysLeuSerSerArgAlaAsnSer-AsnLeuThrLeuThrAspThrGluHisG1	160
Db	677	AGTCCTGCTTGTTCGAGCGGGCCAACTCCAAACCTCACCCCTCACCGACACGGAGCACGA	736
QY	160	uAsnThrGluThr	164
Db	737	GAATACCGAAACT	749
RESULT 6			
AK047423			
LOCUS			
DEFINITION	AK047423	3447 bp mRNA linear HTC 20-SEP-2003	
		Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length	
		enriched library, clone:B930059M16 product:odd Oz/ten-m homolog 1	
		(Drosophila), full insert sequence.	
ACCESSION	AK047423		
VERSION	AK047423.1	GI:26338765	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,		
TITLE	Itch,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to		
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes		
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)		
REFERENCE	2		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,		
	Itch,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.		
	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		
	Genome Res. 10 (10), 1617-1630 (2000)		
JOURNAL	20499374		
MEDLINE	11042159		
PUBMED			
REFERENCE	3		
AUTHORS	Shibata,K., Itch,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,		
	Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,		
	Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,		
	Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,		
	Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,		
	Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,		
	Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.		
	RIKEN integrated sequence analysis (RISA) system--384-format		

JOURNAL	sequencing pipeline with 384 multicapillary sequencer
MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)
PUBMED	20530913
REFERENCE	11076861
AUTHORS	4
	The RIKEN Genome Exploration Research Group Phase II Team and the
	FANTOM Consortium.
	Functional annotation of a full-length mouse cDNA collection
	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research
	Group Phase I & II Team.
	Analysis of the mouse transcriptome based on functional annotation
	of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6
AUTHORS	(bases 1 to 3447)
	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
	Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
	Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
	Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
	Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
	Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
	Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
	Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
	Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
	Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
	Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
	Muramatsu,M. and Hayashizaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
	Physical and Chemical Research (RIKEN), Laboratory for Genome
	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
	Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
	URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
	Fax:81-45-503-9216)
COMMENT	CDNA library was prepared and sequenced in Mouse Genome
	Encyclopedia Project of Genome Exploration Research Group in Riken
	Genomic Sciences Center and Genome Science Laboratory in RIKEN.
	Division of Experimental Animal Research in Riken contributed to
	prepare mouse tissues.
	Please visit our web site for further details.
	URL:http://genome.gsc.riken.go.jp/
	URL:http://fantom.gsc.riken.go.jp/.
FEATURES	Location/Qualifiers
source	1..3447
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM DB:B930059M16"
	/db_xref="MGI:2412967"
	/db_xref="taxon:10090"
	/clone="B930059M16"
	/tissue type="cerebellum"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="10 days neonate"
	862..2472
	/note="unnamed protein product; odd Oz/ten-m homolog 1
	(Drosophila) (MGD MGI:1345185, GB NM_011855, evidence: 1
	BLASTN, 99%, match=3037)
	putative"
	/codon_start=1
	/protein_id="BAC33054.1"
	/db_xref="GI:26338766"
	/translation="MEQTDCKPYQPLSKVKHEMDLAYTSSSEDESDGRKPRQSFNSRE
	TLHEYNOELRRYNSQSRKDVKSTQIEFCETPTLCGYHTDMHSVSRHGQYLE
	MGSPVDTETEGAASPDHALRMIRGMKSEHSSCLSRANSALSLTDDHERKSDGENG
	FKFSPVCCDMEAPADSQDMQSSPHNQFTFRPLPPPPPPHACTCARPPPTVDSLQR
	RSMTTRSQSPSPAAPPTSTQDSVHLHNSVLNSNIPLETRHFLFKHGSSSAIFSA
	SNYPLTSTNTVYSPPPPLRSTFRPAFTFNKPYRCCNWKCTALSAITVTLLLL
	AYVIVHLFGLTWQLQPVGQIYANGISNGNPGTESMDTTSPIGGRVSDKSEKKVQKG
	RAIDTGEVDIGAQMOTIPGGLFWRFQITIHHPILYKFNISLAKDSLGLGIYGRNIPP
	THTFDFVKLMDGKQLVKQDSKSSDDIQHSRNLILTSLQETGTFIEYMQQGPWYAFY

CDS

NDGKKMEQVFLVLTAAIGKLPCLLFFFTV"									
ORIGIN									
Alignment Scores:									
ed. No.:	9.89e-42	Length:	3447						
ore:	720.50	Matches:	169						
cent Similarity:	52.77%	Conservative:	50						
st Local Similarity:	40.72%	Mismatches:	101						
ery Match:	33.79%	Indels:	95						
:	11	Gaps:	15						
Q14-X-AT-28-64-76 (1-400) x AK047423 (1-3447)									
1	MetAspValLysGluArgLysProTyrArgSerLeuThrArg	---	ArgArgAspAlaGlu	19					
862	ATGGAGCAAAACAGACTGCAAACTTATCAGCCTCTGTCCAAAGTCAAGCATGAATGGAT	921							
20	ArgArgTyrThrSerSerAla**SerGluGluGlyLysAlaProGlnLysSerTyr	39							
922	CTAGCTTATACCAAGTCTTCTGATGAGAGTGAAGATGGGAGAAAACCAAGACAGTCATTC	981							
40	SerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer---	58							
982	AACTCCAGGAAACTCTGCATGAGTATACCAAGAGCTGAGGAGGAATTACAATAGCCAG	1041							
59	-----ArgValLysAspIle***ProGlnGluAlaGlu---	GluPheCysArgThrGly	75						
1042	ACTAGAAAGAGGAAGATGTGGAGAAATCTACTCAAGAGATAGAATCTGTGAA-----	1095							
76	***AsnPheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeu	95							
1096	-----ACACCTCCT-----ACTTTG	1110							
96	-----TyrArgThrAspIle---GlyLeuProGlnCysGlyTyrSerMetGlyAla	111							
1111	TGCTCCGGCTACCACAGACATGCACAGTGTCTTCGACATGGCTACCAGCTGGAGATG	1170							
112	GlySerAspAlaAspMetGluAlaAspThrValLeuSerProGluHisProValArgLeu	131							
1171	GGATCTGATGTAGATACAGAGACAGAAAGAGCTGCATCACCTGACCATGCACATGCA	1230							
132	TrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsn	151							
1231	TGGATAAGAGGCATGAATCAGAACACAGTTCCTGTCTGCCAGTAGGGCCAACTCTGCA	1290							
152	LeuThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeuGln	171							
1291	CTGTCTTGACCGACACTGATCATGAACGG---AAGTCTGATGGGAAAAATGGTTTAAA	1347							
171	-----	171							
1348	TTCTCTCCTGTTTGTGTGACATGGAGGCTCCAGCTGATTGGCTCAAGACATGCAAAGC	1407							
172	-----AsnHisAlaArgLeuArg-----ThrProProProProLeuSerHisAla	186							
1408	AGCCACACAAACCAGTTCACCTTCAGACCCCTCCACCACCCGCCCTCCTCCACATGCC	1467							
187	HisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr	206							
1468	TGCACCTGTGCCAGGAAGCCACCTCCTACAGTGGACTCTCTACAAAGAAGATCAATGACT	1527							
207	ProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProProAla	226							
1528	ACCCGACGCCAGCCAGCCAGCT-----GCTCCTGCTCCTCCA	1566							
227	GlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIlePro	246							
1567	ACCAGCACACAGGATTCAGTTTCATCTGCATAACAGCTGGTCTTGAAACAGTAACATACCA	1626							
247	LeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIle	266							
1627	CTGGAGACCAGG-----	1638							

QY	267	GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeu	286
Db	1639	-----CAATTCCTG	1647
QY	287	PheLysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu	305
Db	1648	TTCAAACATGGATCTGGTCTTCTGCTATCTTCAGTCAGCCAGTCAGAACTACCTCTG	1707
QY	306	ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg	325
Db	1708	ACATCTAATACTGTCTACTACCAACCAACCCAGGCCGCTCGAAGCACCTTTTCCCGA	1767
QY	326	ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer	345
Db	1768	CCTGCCTTCACCTTTTAAACAAACCATACAGATGCTGCAATTGGAAGTCACAGCCTTGAGC	1827
QY	346	AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu	365
Db	1828	GCCACTGCAATCACAGTGACTTGGCCTTGTACTAGCCTAT---GTAATTGTACACTTG	1884
QY	366	PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380
Db	1885	TTTGGCCTGACGTGGCAGTGTGCAACCAAGTT---GGACAGATCTAT	1926
RESULT 7			
BQ735622	884 bp	mRNA	linear EST 16-JUL-2002
LOCUS	AGENCOURT 8097117	NICHD_XGC_Emb4	Xenopus laevis cDNA clone
DEFINITION	IMAGE:5542635	5', mRNA	sequence.
ACCESSION	BQ735622	1	GI:21874519
VERSION	BQ735622.1		
KEYWORDS	EST.		
SOURCE	Xenopus laevis	(African clawed frog)	
ORGANISM	Xenopus laevis		
REFERENCE	1	(bases 1 to 884)	
AUTHORS	NCI-CCAP	http://www.ncbi.nlm.nih.gov/ncicgap.	
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Igor Dawid cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Plate: LLAM12242 row: e column: 04 High quality sequence stop: 373.		
FEATURES	Location/Qualifiers		
source	1. .884		
	/organism="Xenopus laevis"		
	/mol_type="mRNA"		
	/db_xref="taxon:8355"		
	/clone="IMAGE:5542635"		
	/dev_stage="embryo, stage 31-32"		
	/lab_host="DH10B (phage-resistant)"		
	/clone_lib="NICHD XGC Emb4"		
	/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1; NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."		
ORIGIN			
Alignment Scores:	5.24e-41	Length:	884
Pred. No.:	699.00	Matches:	160
Score:	60.76%	Conservative:	15

267	GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeu	286		-----CATTCCTG	1647
1639					
287	PheLysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu	305			
1648	TTCAAACATGATCTGGTTCTTCTGCTATCTTCAGTCAGCCAGTCAGAACTACCTCTG	1707			
306	ThrSerSerThrValTyrSerProProArgProLeuProArgSerThrPheAlaArg	325			
1708	ACATCTAATACTGTCTACTCACCAACCACCCAGCCGCTCGAAGCACCTTTTCCCGA	1767			
326	ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer	345			
1768	CCTGCCTTCACTTTTAAACAACCATACAGATGCTGCAATTGGAAGTCACAGCCTTGAGC	1827			
346	AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu	365			
1828	GCCACTGCAATCACAGTGACTTTGGCCTTGTACTAGCCTAT---GTAATTGTACACTTG	1884			
366	PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr	380			
1885	TTTGGCCTGACGTGGCAGTTGCAACCAGTT---GGACAGATCTAT	1926			
RESULT 7					
BQ735622					
LOCUS					
DEFINITION	AGENCOURT_8097117_NICHD_XGC_Emb4_Xenopus laevis cDNA clone				
ACCESSION	BQ735622				
VERSION	BQ735622.1 GI:21874519				
KEYWORDS	EST.				
SOURCE	Xenopus laevis (African clawed frog)				
ORGANISM	Xenopus laevis				
REFERENCE	1 (bases 1 to 884)				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
JOURNAL	Unpublished (1997)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-re@mail.nih.gov Tissue Procurement: Dr. Igor Dawid cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Plate: LLAM1242 row: e column: 04 High quality sequence stop: 373.				
FEATURES	Location/Qualifiers				
source	1..884				
	/organism="Xenopus laevis"				
	/mol_type="mRNA"				
	/db_xref="taxon:8355"				
	/clone="IMAGE:5542635"				
	/dev_stage="embryo, stage 31-32"				
	/lab_host="DH10B (phage-resistant)"				
	/clone_lib="NICHD XGC_Emb4"				
	/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."				
ORIGIN					
Alignment Scores:	5.24e-41	Length:	884		
Pred. No.:	699.00	Matches:	160		
Score:	60.76%	Conservative:	15		

Best Local Similarity: 55.56% Mismatches: 44
Query Match: 32.79% Indels: 69
DB: 13 Gaps: 8

SEQ14-X-AT-28-64-76 (1-400) x BQ735622 (1-884)

```
QY 113 SerAspAlaAspMetGluAlaAspThrValLeuSerProGluHisProValArgLeuTrp 132
Db 26 AACGATTACAGACTCTGAAGGAGACTTTGTC-----55
QY 133 GlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeu 152
Db 56 -----CCTAATTCCTTCTTAGTTAAAGTGATTCGCAAACCTG 94
QY 153 ThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsn 172
Db 95 GCTGTC-----CCTGCTAACGATCAGCGCTGGTTCTGCAGAAAT 133
QY 173 HisAlaArgLeuArgThrProProProLeuSerHisAlaHisThrProAsnGlnHis 192
Db 134 CATTTGAGACTACGAACGCCGCCACGCCCTTAAGCCACCCGCAAAACCCC-----CAC 187
QY 193 HisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSer 212
Db 188 CATGGGCGATCCATCAATTCTTTGAACGGGGCAATTACACCCACGTCAGCAACCCCTAGC 247
QY 213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro 232
Db 248 CCGCGCCCCACGACCACTCTCTGTGGGTGAGCCCCCG---GGCGGTGGCCAGGAGTCC 304
QY 233 AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu 252
Db 305 GTCCAT-----GACAATTGGCTGTGAACAGCAACAATTCCCGCTGGAGACCAGG----- 352
QY 253 GlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly 272
Db 352 -----352
QY 273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysProGlyGlyThr 292
Db 353 -----CATTTCTGTTTAAAGCCTGGATGGACC 379
QY 293 SerProLeuPheCysThrThrSerProGlyTyrProLeuThrThrSerThrValTyrSer 312
Db 380 TCCCTCTGTTCTGCACGACGTCTCCTGGGAACCCCTGACATCCAGCACTGTGTACTCC 439
QY 313 ProProProArgProLeuProArgSerThrPheAlaArg-ProAlaPheAsnLeuLysLy 332
Db 440 CCTCCTCCGAGGACGCTTCCTCGCAGCACTTTCTCGGGGTGGGCTTTAAACCTTAAGAA 499
QY 332 sProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAlaTh 352
Db 500 GCCGTACAAATACTGCAACTGGAAATGTGCCGCTGAGCGCCATCATCATCTCCGTCCAC 559
QY 352 rLeuValIleLeuLeuAla-TyrPheValAlaMetHisLeu--PheGlyLeuAsnTrp-- 370
Db 560 TCTGAAGATCCTCTGGGATTATTTTCATCGCCATGCACCTTGTGTCGCTAAACTGGGC 619
QY 371 -HisLeuGlnProMetGluGlyGlnMetTyrGlu-IleThrGlu---AspThrAlaSerS 389
Db 620 CCCTGCTGCCGAACGAAGGGCCAGAAGTATGACACTCCCGAAGGACAGAGGGCGGCAA 679
QY 389 erTrpProValPro 393
Db 680 TTTGGCCACGTCCC 693
```

RESULT 8
CB544750
LOCUS
DEFINITION
CB544750
AMGNNUC:TRYPI-00009-A5-A trypl (10582) Rattus norvegicus cDNA clone
trypl-00009-a5 5', mRNA sequence.
ACCESSION
CB544750
VERSION
CB544750.1 GI:29428647

CB544750
LOCUS
DEFINITION
603760992F1 CSEQRBN21 Gallus gallus cdna clone ChEST679g24 5', mRNA
sequence.
BU474512
ACCESSION
BU474512
VERSION
BU474512.1 GI:25968089
KEYWORDS
EST.

KEYWORDS
SOURCE
ORGANISM

EST.
Rattus norvegicus (Norway rat)
Rattus norvegicus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 402)
Amgen EST Program.
Amgen Rat EST Program
Unpublished (2003)
Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00009 row: a column: 5.
Location/Qualifiers
1. 402
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="trypl-00009-a5"
/tissue_type="prostate"
/clone_lib="trypl (10582)"
/note="Vector: pYYA-4L; Site_1: HindIII; Site_2: NotI;
rat prostate"

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 1.79e-40 Length: 402
Score: 685.00 Matches: 126
Percent Similarity: 99.22% Conservative: 2
Best Local Similarity: 97.67% Mismatches: 1
Query Match: 32.13% Indels: 0
DB: 14 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x CB544750 (1-402)

```
QY 250 ArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAsp 269
Db 2 AGAAACCTAGGCAAGCAGCATTCTTAGGACATTGCAGGACAACTCATTTGAGATGGAC 61
QY 270 IleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeuPheLysPro 289
Db 62 ATTCTCAGCGCCTCCGCCATGATGGGGCTTACAGTGACGGGCACTTCCTCTTCAAGCCC 121
QY 290 GlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThr 309
Db 122 GGAGGCACCTCCCCACTCTTCTGCACCACTCCCGAGGTACCCCTAACGTCCTAGCACC 181
QY 310 ValTyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsn 329
Db 182 GTGTATTGCGCCCAACCCCGGCCCTGCCCCGAGGCTTTTCGAGACCAGCCTTTAAC 241
QY 330 LeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIle 349
Db 242 CTCAAGAAGCCTTCCAAGTACTGCAACTGGAAGTGTGGGCACCTAAGCGCCATCCTCATC 301
QY 350 SerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaAlaMetHisLeuPheGlyLeuAsn 369
Db 302 TCAGCTACGCTCGTCATCTGCTGTCGATACTTTGTGGCCATGCACCTGTTGGCCTAAAC 361
QY 370 TrpHisLeuGlnProMetGluGlyGln 378
Db 362 TGGCACCTGCAGCCGATGGAGGGGCAG 388
```

RESULT 9
BU474512
LOCUS
DEFINITION
603760992F1 CSEQRBN21 Gallus gallus cdna clone ChEST679g24 5', mRNA
sequence.
BU474512
ACCESSION
BU474512
VERSION
BU474512.1 GI:25968089
KEYWORDS
EST.

SOURCE
ORGANISM
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
AUTHORS
1 (bases 1 to 789)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE
PUBMED
22335534
12445392
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .789
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="ChEST679g24"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBN21"
/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 1.09e-39 Length: 789
Score: 680.00 Matches: 137
Percent Similarity: 89.70% Conservative: 11
Best Local Similarity: 83.03% Mismatches: 16
Query Match: 31.89% Indels: 3
DB: 13 Gaps: 0

SEQ14-X-AT-28-64-76 (1-400) x BU474512 (1-789)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
|||||
Db 256 ATGGATGTAAAGAAAGGAAACCGTATCGATCTCTGACTCGGCGCCGCGACACGGAGCGC 315
|||||

QY 21 ArgTyrThrSerSerAla**SerGluGluGlyLysAlaproGlnLysSerTyrSer 40
|||||
Db 316 CGCTACACCAAGCTCTTCAGCCGAGAGTGAGCAGCAGCAAGGCTCCTCAGAAGTCTATAGC 375
|||||

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||||
Db 376 TCCAGTGAGACCCCTGAAGGCTTATGATCAAGACTCCAGGTTGACCTACAGCAATCGGGTC 435
|||||

QY 61 LysAspIle**ProGlnGluAlaGluPheCysArgThrGly***AsnPheThrLeu 80
|||||
Db 436 AAAGACATGCTGCACCAGGAGGCTGATGAATCTGTGCCGAGCAGGAGCCAACTTCTCTTTG 495
|||||

QY 81 ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
|||||
Db 496 CGGGAGCTGGGTCTTGAAGATGTGACTCCACCCATGGGACTTTGTACCGGACTGATATT 555
|||||

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
|||||
Db 556 GGGCTGCCTCACTGTGGTACTCCATCAGCGCTGGCTCAGATGCCGACACGGAGCGGAT 615
|||||

QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
|||||
Db 616 GTGGTCATGTACCTGAGCATCTGTGAGGCTCTGGGACGCAACACCAATCCGACGC 675
|||||

QY 141 SerSer-CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1 160
|||||
Db 676 AGCTCCCTGCTTGTGAGCCGGGC-AACTCCACCTCACCTC-ACCGACACGGAGCAGCA 733
|||||

QY 160 uAsnThrGluThr 164
|||||
Db 734 GAATACCGAAACT 746
|||||

RESULT 10
BU119163
LOCUS
DEFINITION
BUI19163 CSEQCHL16 Gallus gallus cDNA clone ChEST137111 5', mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE
AUTHORS
1 (bases 1 to 887)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
COMMENT
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .887
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 15I"
/db_xref="taxon:9031"
/clone="ChEST137111"
/sex="Female"
/tissue_type="not cerebrum or cerebellum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHL16"
/note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
[Stratagene] vector to accommodate cDNA produced with the
T-trimmed protocol (Construction of uni-directionally
cloned cDNA libraries from messenger RNA for improved 3'
end DNA sequencing by Glenn Fu, et al. U.S. Patent
6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
Ligate in double stranded adaptor containing BsgI and
BamHI sites [5'ggcgcgtgcagcccgatccgagcaaaaaag]
[5'aattcttttttcggtcgatccgggctgcacgc]"

FEATURES
source
ORIGIN
Alignment Scores:

Pred. No.:	2.75e-39	Length:	887
Score:	675.50	Matches:	144
Percent Similarity:	85.79%	Conservative:	13
Best Local Similarity:	78.69%	Mismatches:	24
Query Match:	31.68%	Indels:	5
DB:	13	Gaps:	0
SEQ14-X-AT-28-64-76 (1-400) x BU119163 (1-887)			
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg	20
Db	343	ATGGATGTAAGAAAGAAACCGTATCGATCTCTGACTCGGCGCCGACACGGAGCGC	402
QY	21	ArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaProGlnLysSerTyrSer	40
Db	403	CGCTACACCAGCTCTTCAGCCGAGAGTGAGGACAGCAAGGCTCCTCAGAAGTCCTATAGC	462
QY	41	SerSerGluThrLeuLysAlaTyrzAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60
Db	463	TCCAGTGAGACCCCTGAAGGCTTATGATCAAGACTCCAGGTTGACCTACAGCAATCGGGTC	522
QY	61	LysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeu	80
Db	523	AAAGACATGGTGCACACGAGGCTGATGAATTCTGCCGAGCAGGAGCCAATTCTCTTTG	582
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100
Db	583	CGGAGTGGGTCTTGAAGATGTGACTCCCAACCCATGGGACTTGTACCGGACTGATATT	642
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120
Db	643	GGGTGCCTCACTGTGGCTACTCCATCAGCGCTGGCTCAGATGCCGACACGGAGGCGGAT	702
QY	121	ThrValLeuSer-ProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyAr	140
Db	703	GTGGTCATGTCAACCTGAGCATCTCTGTGAGGCTCTGGGGAGCGCAACCAATCCGGACG	762
QY	140	gSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGl	160
Db	763	CAGTCTCTGCTTGTGAGCGCGGCCAACTCCAATC-ACCCTCACCGACACGGAAGCCCG	821
QY	160	uAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg-ThrProp	180
Db	822	AGAATACGAAAT-GATCATCTCCATATCTTCA-AATCATTCAGAGACTCCGAATCCACAC	879
QY	180	roPro 181	
Db	880	CTCCA 884	
RESULT 11			
AK034286			
LOCUS			
DEFINITION	AK034286	4556 bp	mRNA linear HTC 18-SEP-2003
	Mus musculus adult male diencephalon cDNA, RIKEN full-length		enriched library, clone:9330172M18 product:odd Oz/ten-m homolog 1
	(Drosophila), full insert sequence.		
ACCESSION	AK034286		
VERSION	AK034286.1	GI:26083876	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
	Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to		
	prepare full-length cDNA libraries for rapid discovery of new genes		

JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
REFERENCE	3
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 4556)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/.
FEATURES	Location/Qualifiers
source	1. .4556
	/organism="Mus musculus"
	/mol_type="mRNA"
	/strain="C57BL/6J"
	/db_xref="FANTOM_DB:9330172M18"
	/db_xref="MGI:2398485"
	/db_xref="taxon:10090"
	/clone="9330172M18"
	/sex="male"
	/tissue_type="diencephalon"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
	1. .4556
	/note="odd Oz/ten-m homolog 1 (Drosophila)
	(MGD MGI:1345185, GB NM_011855, evidence: BLASTN, 99%, match=3037)"
misc_feature	1. .4556


```
SEQ14-X-AT-28-64-76 (1-400) x AL9222332 (1-490)
QY 3 VallysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArgArgTyr 22
Db 489 GTGAAGGAACGCAGACCTACCGCTCTCTGACCTCCAGGCGGACACGAGCGCGCTAC 430
QY 23 ThrSerSerSerAla***SerGluGluGluPheCysArgThrGly***AsnPheThrLeuArgGlu 82
Db 429 ACCAGCTCATCTGCCGACAGCGAGGATGGCAAGATCAACCCCTAAATCTTACAGCTCGAGT 370
QY 43 GluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAsp 62
Db 369 GAGACCCTCAAGCCTTCGACCAGGACTCCAGACTGGCTTATGGCAGCGCGTCAAAGAC 310
QY 63 Ile***ProGlnGluAlaGluGluPheCysArgThrGly***AsnPheThrLeuArgGlu 82
Db 309 CTGGTGCAACCATGAGGCCGACGAGTTTCAGCAGACAAGGCGCGACTTTTCTCTCAGAGAC 250
QY 83 LeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeu 102
Db 249 ATGGCCTTCGAGATCCCGTGCCGCCACATATGGCAGCGTACCGGACAGAAATGGGCCTT 190
QY 103 ProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrVal 122
Db 189 CCCCACCGGCACTACTCAGTGAGTGTGGCATCAGACGCTGACACAGAAACAGACGGCATA 130
QY 123 LeuSerProGluHisProValArgLeuTrpGlyArgSer---ThrArgSerGlyArgSer 141
Db 129 ATGTCCCGGAGCACGCGCTCGGCCTCTGGGCGCGCAGCAACACCAAAATCCGGCGCGAGT 70
QY 142 SerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsn 161
Db 69 TCCTGCTGTCCAGCAGAGCCCACTCAACCTTACGCTCACCGACACCGAGCATGAAAC 10
QY 162 ThrGlu 163
Db 9 ACTGAG 4
RESULT 13
BG036207
LOCUS 602326960F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428351 5',
DEFINITION mRNA sequence.
ACCESSION BG036207
VERSION BG036207.1 GI:12431132
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10178 row: h column: 16
High quality sequence stop: 713.
Location/Qualifiers
1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4428351"
/tissue_type="adenocarcinoma, cell line"
FEATURES
source
```

```
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sall; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
ORIGIN
Alignment Scores:
Pred. No.: 1.95e-31 Length: 870
Score: 568.00 Matches: 109
Percent Similarity: 99.09% Conservative: 0
Best Local Similarity: 99.09% Mismatches: 1
Query Match: 26.64% Indels: 1
DB: 10 Gaps: 0
SEQ14-X-AT-28-64-76 (1-400) x BG036207 (1-870)
QY 291 GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal 310
Db 2 GGCACCTCCCGCTCTTCTGCACCAATCACCAGG-TACCCACTGACGTCCAGCACAGTG 60
QY 311 TyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu 330
Db 61 TACTCTCTCGCCCCGACCCCTGCCCGCAGCACCTTCGCCCGCGCGCCTTTAACCTC 120
QY 331 LysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSer 350
Db 121 AAGAAGCCCTCCAAGTACTGTAAGTGAAGTGGCAGCCCTGAGCGCCATCGTCATCTCA 180
QY 351 AlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTrp 370
Db 181 GCCACTCTGGTCATCTCTGCTGGCATACTTTGTGGCCATGCACCTGTTGGCCTAAACTGG 240
QY 371 HisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTrp 390
Db 241 CACCTGCAGCCGATGGAGGGGAGATGTATGATCATCCGAGGACACAGCCAGCAGTTGG 300
QY 391 ProValProThrAspValSerLeuTyrPro 400
Db 301 CCTGTGCCAACCGACGCTCTCCCTATACCCC 330
RESULT 14
AY405291
LOCUS Homo sapiens HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405291
VERSION AY405291.1 GI:39761265
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 511)
AUTHORS Clark,A.G., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 511)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
```


FEATURES source them based on alignment. Location/Qualifiers 1..511 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>511 /locus_tag="HCM2175" ORIGIN Alignment Scores: 3.19e-28 Length: 511 Pred. No.: 519.50 Matches: 106 Score: 78.66% Conservative: 23 Percent Similarity: 64.63% Mismatches: 32 Best Local Similarity: 24.37% Indels: 3 Query Match: 29 Gaps: 3 SEQ14-X-AT-28-64-76 (1-400) x AY405291 (1-511) QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg--ArgArgAspAlaGlu 19 Db 1 ATGGATGTGAAGAACGCGAGGCCCTTACTGCTCCCTGACCAAGACGACGAGAGAGGAA 60 QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38 Db 61 CGGCGGTACACAAATTCTCCGACAGCAATGAGGAGTGCCGGGTACCCACACAGAAGTCC 120 QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57 Db 121 TACAGTTCACGCGAGACATTGAAAGCTTTTGATCATGATTCTCCGGGTGCTTTACGGC 180 QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77 Db 181 AACAGAGTGAAGGATTGGTTTCACAGACAAGCAGACGAGTTCACTAGACAAGGACAGAAT 240 QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97 Db 241 TTTACCTTAAGCAGATTAGGAGTTTGTGAACAGCAACTCGAAGAGGACTGGCATTTTGT 300 QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117 Db 301 GCGGAAATGGGGCTCCCTCACAGAGGTACTCTATCAGTGCAGGGTCAGATGCTGATACT 360 QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137 Db 361 GAAATGAAGCAGTATGTCCCACAGAGCATGCCATGAGACTTTGGGGCAGGGGTCAA 420 QY 138 SerGlyArgSerSerCysLeuSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157 Db 421 TCAGGCGCAGCTCCTGCTGTCAAGTCGGTCCAACTCAGCCCTCACCCCTGACAGATACG 480 QY 158 GluHisGluAsn 161 Db 481 GAGCACGAAAC 492 RESULT 15 AY405292 511 bp DNA linear GSS 12-DEC-2003 LOCUS Pan troglodytes HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence, DEFINITION genomic survey sequence. ACCESSION AY405292 VERSION AY405292.1 GI:39761266 KEYWORDS GSS. SOURCE Pan troglodytes (chimpanzee) ORGANISM Pan troglodytes BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; MAMMALIA; Eutheria; Primates; Catarrhini; Homnidae; Pan. REFERENCE 1 (bases 1 to 511) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios JOURNAL Science 302 (5652), 1960-1963 (2003) PUBMED 14671302 REFERENCE 2 (bases 1 to 511) AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. TITLE Direct Submission JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. FEATURES Location/Qualifiers source 1..511 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>511 /locus_tag="HCM2175" ORIGIN Alignment Scores: 3.19e-28 Length: 511 Pred. No.: 519.50 Matches: 106 Score: 78.66% Conservative: 23 Percent Similarity: 64.63% Mismatches: 32 Best Local Similarity: 24.37% Indels: 3 Query Match: 29 Gaps: 3 SEQ14-X-AT-28-64-76 (1-400) x AY405292 (1-511) QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg--ArgArgAspAlaGlu 19 Db 1 ATGGATGTGAAGAACGCGAGGCCCTTACTGCTCCCTGACCAAGACGACGAGAGGAA 60 QY 20 ArgArgTyrThrSerSerSerAla***SerGluGluGlyLysAlaPro---GlnLysSer 38 Db 61 CGGCGGTACACAAATTCTCCGACAGCAATGAGGAGTGCCGGGTACCCACACAGAAGTCC 120 QY 39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57 Db 121 TACAGTTCACGCGAGACATTGAAAGCTTTTGATCATGATTCTCCGGGTGCTTTACGGC 180 QY 58 SerArgValLysAspIle***ProGlnGluAlaGluGluPheCysArgThrGly***Asn 77 Db 181 AACAGAGTGAAGGATTGGTTTCACAGACAAGCAGACGAGTTCACTAGACAAGGACAGAAT 240 QY 78 PheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArg 97 Db 241 TTTACCTTAAGCAGATTAGGAGTTTGTGAACAGCAACTCGAAGAGGACTGGCATTTTGT 300 QY 98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117 Db 301 GCGGAAATGGGGCTCCCTCACAGAGGTACTCTATCAGTGCAGGGTCAGATGCTGATACT 360 QY 118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArg 137 Db 361 GAAATGAAGCAGTATGTCCCACAGAGCATGCCATGAGACTTTGGGGCAGGGGTCAA 420 QY 138 SerGlyArgSerSerCysLeuSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157 Db 421 TCAGGCGCAGCTCCTGCTGTCAAGTCGGTCCAACTCAGCCCTCACCCCTGACAGATACG 480 QY 158 GluHisGluAsn 161 Db 481 GAGCACGAAAC 492 Search completed: August 15, 2004, 00:43:42 Job time : 3356 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 152.319 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_2650_2725
Perfect score: 391
Sequence: 1 SQINTVLNGTRTRYTDIQLQ.....EQRLREGEGGLRAWTEGEK 76

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn_1_1_2156@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
*					

1	391	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
2	391	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
3	391	100.0	8355	13	US-10-383-201-55	Sequence 55, Appl
4	391	100.0	8438	13	US-10-042-865-1	Sequence 1, Appli
5	308	78.8	12880	16	US-10-295-027-927	Sequence 927, App
6	296	75.7	3614	13	US-10-342-887-1743	Sequence 1743, Ap
7	296	75.7	3614	13	US-10-172-118-1743	Sequence 1743, Ap
8	296	75.7	8473	17	US-10-038-854-39	Sequence 39, Appl
9	296	75.7	8487	17	US-10-038-854-41	Sequence 41, Appl
10	296	75.7	8645	17	US-10-038-854-37	Sequence 37, Appl
11	296	75.7	8675	17	US-10-038-854-35	Sequence 35, Appl
12	249.5	63.8	8409	9	US-09-808-602-79	Sequence 79, Appl
13	249.5	63.8	8409	10	US-09-800-198-67	Sequence 67, Appl
14	246.5	63.0	3217	15	US-10-198-846-13976	Sequence 13976, A
15	246.5	63.0	6560	9	US-09-808-602-76	Sequence 76, Appl
16	246.5	63.0	6560	10	US-09-800-198-64	Sequence 64, Appl
17	246.5	63.0	8575	13	US-10-072-012-143	Sequence 143, App
18	246.5	63.0	8689	9	US-09-808-602-78	Sequence 78, Appl
19	246.5	63.0	8689	10	US-09-800-198-66	Sequence 66, Appl
20	246.5	63.0	9058	16	US-10-144-194A-79	Sequence 79, Appl
21	246.5	63.0	9695	16	US-10-144-194A-81	Sequence 81, Appl
22	246.5	63.0	9729	9	US-09-808-602-12	Sequence 12, Appl
23	246.5	63.0	9729	10	US-09-800-198-12	Sequence 12, Appl
24	246.5	63.0	9826	9	US-09-808-602-7	Sequence 7, Appli
25	246.5	63.0	9826	10	US-09-800-198-7	Sequence 7, Appli
26	239.5	61.3	8797	9	US-09-808-602-74	Sequence 74, Appl
27	239.5	61.3	8797	9	US-09-808-602-77	Sequence 77, Appl
28	239.5	61.3	8797	10	US-09-800-198-62	Sequence 62, Appl
29	239.5	61.3	8797	10	US-09-800-198-65	Sequence 65, Appl
30	117.5	30.1	464	10	US-09-918-995-27921	Sequence 27921, A
31	90	23.0	65	10	US-09-908-975-28311	Sequence 28311, A
32	77.5	19.8	1078	9	US-09-840-787-90	Sequence 90, Appl
33	77.5	19.8	2015	10	US-09-919-039-262	Sequence 262, App
34	75.5	19.3	992	13	US-09-823-245A-447	Sequence 447, App
35	75.5	19.3	1171	15	US-10-198-846-10358	Sequence 10358, A
36	75	19.2	4513	13	US-10-240-425-1459	Sequence 1459, Ap
37	75	19.2	6457	9	US-09-880-107-3389	Sequence 3389, Ap
38	74.5	19.1	2940917	13	US-10-027-632-174763	Sequence 174763,
39	74.5	19.1	2940917	16	US-10-027-632-174763	Sequence 174763,
40	74	18.9	591	17	US-10-021-323-785	Sequence 785, App
41	73.5	18.8	285	9	US-09-969-708-147	Sequence 147, App
42	73.5	18.8	445	14	US-10-001-835-20	Sequence 20, Appl
43	73.5	18.8	791	9	US-09-822-849A-578	Sequence 578, App
44	73.5	18.8	814	13	US-10-424-599-46208	Sequence 46208, A
45	73	18.7	60	10	US-09-908-975-15821	Sequence 15821, A

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)...(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 3.16e-47 Length: 8354
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-383-201-43 (1-8354)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7982 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAG 8041

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8042 TACGGGGCACTGTGCTTGAACACACGCTACGGGCAACGTTGGATGAGGAGGACCG 8101

QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
Db 8102 GTCCTGGAGCTGGCCGGCAGAGCCGTGCCCAAGCGTGGCCCGCGCAGCAGAGA 8161

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
Db 8162 CTGCGGGAAGGGAGGAGGAGGCGCTGCGGGCCTGGACAGAGGGGGAGAAG 8209

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.: 3.16e-47 Length: 8354
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-029-020-13 (1-8354)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7982 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAG 8041

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8042 TACGGGGCACTGTGCTTGAACACACGCTACGGGCAACGTTGGATGAGGAGGACCG 8101

QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
Db 8102 GTCCTGGAGCTGGCCGGCAGAGCCGTGCCCAAGCGTGGCCCGCGCAGCAGAGA 8161

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
Db 8162 CTGCGGGAAGGGAGGAGGAGGCGCTGCGGGCCTGGACAGAGGGGGAGAAG 8209

RESULT 3
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(8325)
US-10-383-201-55

```


Alignment Scores: 3.16e-47 Length: 8355
Pred. No.: 391.00 Matches: 76
Score: 391.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-383-201-55 (1-8355)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 7966 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 8025

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8026 TACGGGCACTGTGCTTGAACACACACGCTACGGGACAACGTTGGATGAGGAGGACCG 8085

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
Db 8086 GTCCTGGAGCTGGCCCGGCAGAGAGCCGTGGCCCAAGCGTGGGCCCGCAGCAGCAGAGA 8145

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
Db 8146 CTGCGGGAAGGGAGGAGGCGCTGCGGGCCTGGACAGAGGGGGAGAG 8193

RESULT 4

US-10-042-865-1

; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876

; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores: 3.2e-47 Length: 8438
Pred. No.: 391.00 Matches: 76
Score: 391.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-042-865-1 (1-8438)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 8033 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAG 8092

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8093 TACGGGCACTGTGCTTGAACACACACGCTACGGGACAACGTTGGATGAGGAGGACCG 8152

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
Db 8153 GTCCITGGAGCTGGCCCGGCAGAGAGCCGTGCGCAAGCGTGGGCCCGCAGCAGCAGAGA 8212

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
Db 8213 CTGCGGGAAGGGAGGAGGCGCTGCGGGCCTGGACAGAGGGGGAGAG 8260

RESULT 5

US-10-295-027-927

; Sequence 927, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10

PRIORITY APPLICATION NUMBER: US 60/355,250		Best Local Similarity: 71.05%		Mismatches: 10	
PRIORITY FILING DATE: 2002-02-08		Query Match: 75.70%		Indels: 0	
PRIORITY APPLICATION NUMBER: US 60/356,714		DB: 13		Gaps: 0	
US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-342-887-1743 (1-3614)					
QY	1	SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20			
DB	2924	TCGCAGTCCACCACCGTGTGTGAACGGCAGGACCGCGACGTTCGCGGACGTGGAGATGCAG 2983			
QY	21	TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40	:::	:::	:::
DB	2984	TTCGGCGCGCTGGCGCTGCACGTGCGCTACGGCATGACCCCTGGACGAGGAGGCGCGC 3043			
QY	41	ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnLeuGlnArg 60	:::	:::	:::
DB	3044	ATCCTGGAGCAGGCGCGCGCGCTCGCCCGGCGCTGGCGCGCGCGAGCAGCAGCGCGC 3103			
QY	61	LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76	:::	:::	:::
DB	3104	GTGCGGACGCGGAGGAGGCGCGCGCTCTGGACGAGGCGCGGAGGCGCGGAGGAGG 3151			
RESULT 7					
US-10-172-118-1743					
; Sequence 1743, Application US/10172118					
; Publication No. US20030224374A1					
; GENERAL INFORMATION:					
; APPLICANT: Dai, Hongyue					
; APPLICANT: He, Yudong					
; APPLICANT: Linsley, Peter					
; APPLICANT: Mao, Mao					
; APPLICANT: Roberts, Chris					
; APPLICANT: Van 't Veer, Laura					
; APPLICANT: Van de Vijver, Marc					
; APPLICANT: Bernards, Rene					
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients					
; FILE REFERENCE: 9301-175-999					
; CURRENT APPLICATION NUMBER: US/10/172,118					
; CURRENT FILING DATE: 2002-06-14					
; PRIOR APPLICATION NUMBER: 60/380,770					
; PRIOR FILING DATE: 2002-05-14					
; NUMBER OF SEQ ID NOS: 2699					
; SEQ ID NO 1743					
; LENGTH: 3614					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
; PUBLICATION INFORMATION:					
; DATABASE ACCESSION NUMBER: NM_018104					
; DATABASE ENTRY DATE: 2001-06-18					
US-10-172-118-1743					
Alignment Scores:		1.64e-33		Length: 3614	
Pred. No.:		296.00		Matches: 54	
Score:		86.84%		Conservative: 12	
Percent Similarity:		71.05%		Mismatches: 10	
Best Local Similarity:		75.70%		Indels: 0	
Query Match:		13		Gaps: 0	
DB:					
US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-172-118-1743 (1-3614)					
QY	1	SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20			
DB	2924	TCGCAGTCCACCACCGTGTGTGAACGGCAGGACCGCGACGTTCGCGGACGTGGAGATGCAG 2983			
QY	21	TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40	:::	:::	:::
DB	2984	TTCGGCGCGCTGGCGCTGCACGTGCGCTACGGCATGACCCCTGGACGAGGAGGCGCGC 3043			
QY	41	ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnLeuGlnArg 60	:::	:::	:::
DB	3044	ATCCTGGAGCAGGCGGCGGAGCGCGCTCGCCCGGCGCTGGCGCGGCGGAGGAGGCGCGC 3103			

PRIORITY APPLICATION NUMBER: US 60/355,250		Best Local Similarity: 71.05%		Mismatches: 10	
PRIORITY FILING DATE: 2002-02-08		Query Match: 75.70%		Indels: 0	
PRIORITY APPLICATION NUMBER: US 60/356,714		DB: 13		Gaps: 0	
US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-342-887-1743 (1-3614)					
QY	1	SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20			
DB	7880	TCCAGATGACTTCTCTGTTGAATGGAGGACTAGACGGTTTCAGATATTCAGCTCCAG 7939			
QY	21	TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40	:::	:::	:::
DB	7940	CATGGAGCCCTGTGCTTCAACATCCGGTATGGACAACACTGTCGAGAGGAAAGAAATCAC 7999			
QY	41	ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnLeuGlnArg 60	:::	:::	:::
DB	8000	GTGTTGGAGATTGCCAGACAGCGCGCAGTAGTGCCCGAGGCCTGGACTAAGCAACAAGAAGG 8059			
QY	61	LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76	:::	:::	:::
DB	8060	CTGCAAGAGGGGAAGAGGGGATTAGGGCATGGACAGAGGGGAAAG 8107			
RESULT 6					
US-10-342-887-1743					
; Sequence 1743, Application US/10342887					
; Publication No. US20040058340A1					
; GENERAL INFORMATION:					
; APPLICANT: Dai, Hongyue					
; APPLICANT: He, Yudong					
; APPLICANT: Linsley, Peter S.					
; APPLICANT: Mao, Mao					
; APPLICANT: Roberts, Christopher J.					
; APPLICANT: Van 't Veer, Laura Johanna					
; APPLICANT: Van de Vijver, Marc J.					
; APPLICANT: Bernards, Rene					
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients					
; FILE REFERENCE: 9301-188-999					
; CURRENT APPLICATION NUMBER: US/10/342,887					
; CURRENT FILING DATE: 2003-01-15					
; PRIOR APPLICATION NUMBER: 60/298,918					
; PRIOR FILING DATE: 2001-06-18					
; PRIOR APPLICATION NUMBER: 60/380,710					
; PRIOR FILING DATE: 2002-05-14					
; PRIOR APPLICATION NUMBER: 10/172,118					
; PRIOR FILING DATE: 2002-06-14					
; NUMBER OF SEQ ID NOS: 2699					
; SEQ ID NO 1743					
; LENGTH: 3614					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-10-342-887-1743					
Alignment Scores:		1.64e-33		Length: 3614	
Pred. No.:		296.00		Matches: 54	
Score:		86.84%		Conservative: 12	

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
:::|||||:::|||||
Db 3104 GTGCGCAGCGGAGGAGCGCGCGCTCTGGACGGAGGCGGAGAAG 3151

RESULT 8

US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-038-854-39

Alignment Scores:
Pred. No.: 4.89e-33 Length: 8473
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12

Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-038-854-39 (1-8473)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||||:::|||||
Db 7782 TCGCAGTCCACACCGTGGTGAACGGCAGCGCAGCGCAGGTTCCGCGGACGTGGAGATGCAG 7841
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
:::|||||:::|||||
Db 7842 TTCGGCGCGCTGGCGCTGCACGTGGCTACGGCATGACCCCTGGACGAGGAGAGGCGCGC 7901
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
:::|||||:::|||||
Db 7902 ATCCTGGAGCAGCGCGCGCGCGCTCGCCCGGCGCTCGCCCGCGCGCAGCAGCGC 7961
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
:::|||||:::|||||
Db 7962 GTGCGCGCAGCGGAGGAGGCGCGCGCTCTGGACGGAGGCGGAGAAG 8009

RESULT 9

US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens


```

; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.:          4.9e-33          Length:      8487
Score:              296.00          Matches:      54
Percent Similarity: 86.84%          Conservative: 12
Best Local Similarity: 71.05%        Mismatches:   10
Query Match:        75.70%          Indels:       0
DB:                  17              Gaps:         0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-038-854-41 (1-8487)

Qy      1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      7778 TCGCAGTCCACACCGGTGGTGAACGCGACGCGCGAGGTTCCGCGACGTGGAGATGCAG 7837

Qy      21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
      ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
Db      7838 TTCGGCGCGCTGGCGCTGCACGTGCTACGCGATGACCCCTGGAGGAGGAGGCGCGC 7897

Qy      41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
      ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
Db      7898 ATCCTGGAGCAGCGCGCGCAGCGCGCTCGCCCGGCTGGCGCGCGAGCAGCAGCGC 7957

Qy      61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
      ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
Db      7958 GTGCGCGACGCGGAGGAGGCGCGCGCTCTGGACGGAGGCGGAGAG 8005

RESULT 10
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenna
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
```

```

; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.:          5.02e-33          Length:      8645
Score:              296.00          Matches:      54
Percent Similarity: 86.84%          Conservative: 12
Best Local Similarity: 71.05%        Mismatches:   10
Query Match:        75.70%          Indels:       0
DB:                  17              Gaps:         0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-038-854-37 (1-8645)

Qy      1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      7954 TCGCAGTCCACACCGGTGGTGAACGCGACGCGCGAGGTTCCGCGACGTGGAGATGCAG 8013

Qy      21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
      ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
Db      8014 TTCGGCGCGCTGGCGCTGCACGTGCTACGCGATGACCCCTGGAGGAGGAGGCGCGC 8073

Qy      41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
      ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
Db      8074 ATCCTGGAGCAGCGCGCGCAGCGCGCTCGCCCGGCTGGCGCGCGAGCAGCAGCGC 8133

Qy      61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
      ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
Db      8134 GTGCGCGACGCGGAGGAGGCGCGCGCTCTGGACGGAGGCGGAGAG 8181

RESULT 11
US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
```



```

; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-800-198-67

Alignment Scores:
Pred. No.: 4.24e-26 Length: 8409
Score: 249.50 Matches: 48
Percent Similarity: 75.95% Conservative: 12
Best Local Similarity: 60.76% Mismatches: 16
Query Match: 63.81% Indels: 3
DB: 10 Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-800-198-67 (1-8409)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 8038 TCCCAGCCAAACCTCTTATCAACGGAAGGACTCGACGGTTACAAACATCGAGTTTCAG 8097

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 8098 TATCCACCCTGCTGATCAACATCCGCTACGGGCTACCGCCGACACGCTGGATGAGGAG 8157

QY 38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlu 57
Db 8158 AAGGCAGAGTGCTAGACAGGTCGCGCAGCGAGCCCTGGGGTCGCGCTGGGCCAAAGAG 8217

QY 58 GlnGlnArgLeuArgGluGluGlyGluArgAlaTrpThrGluGlyGluLys 76
Db 8218 CAGCAGAAGGCACGGGATGGCCGCGAGGGCAGCCGCGTATGGACAGACGGAGAAG 8274

RESULT 14
US-10-198-846-13976/c
; Sequence 13976, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13976
; LENGTH: 3217
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-13976

Alignment Scores:
Pred. No.: 3.46e-26 Length: 3217
Score: 246.50 Matches: 49
Percent Similarity: 75.95% Conservative: 11
Best Local Similarity: 62.03% Mismatches: 16
Query Match: 63.04% Indels: 3
DB: 15 Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-10-198-846-13976 (1-3217)
```

```

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 1717 TCCCAGCCACGCTGCTCAACGGCAGGACTCGAAGGTTACGAACATTGAGTCCAG 1658

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 1657 TACTCCACGCTGCTGCTCAGCATCCGCTATGGCCTACCCCCGACACCCCTGGACGAAGAG 1598

QY 38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlu 57
Db 1597 AAGGCCCGCGTCTCTGGACAGGCGAGACAGAGGGCCCTGGGCACGGCCTGGGCCAAGAG 1538

QY 58 GlnGlnArgLeuArgGluGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 1537 CACGAGAAAGCCAGGACGGGAGAGAGGGGAGCCGCTGTGGACTGAGGGCGAGAAG 1481

RESULT 15
US-09-808-602-76
; Sequence 76, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herzman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 6560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-808-602-76

Alignment Scores:
Pred. No.: 8.64e-26 Length: 6560
Score: 246.50 Matches: 49
Percent Similarity: 75.95% Conservative: 11
Best Local Similarity: 62.03% Mismatches: 16
Query Match: 63.04% Indels: 3
DB: 9 Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-808-602-76 (1-6560)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 4844 TCCCAGCCACGCTGCTGGTCAACGGCAGGACTCGAAGGTTACGAACATTGAGTCCAG 4903

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db 4904 TACTCCACGCTGCTGCTCAGCATCCGCTATGGCCTACCCCCGACACCCCTGGACGAAGAG 4963

QY 38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlu 57
Db 4964 AAGGCCCGCGTCTCTGGACCCAGGCGAGACAGAGGGCCCTGGGCACGGCCTGGGCCAAGAG 5023

QY 58 GlnGlnArgLeuArgGluGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 5024 CAGCAGAAAGCCAGGACGGGAGAGAGGGGAGCCGCTGTGGACTGAGGGCGAGAAG 5080

Search completed: August 14, 2004, 19:23:12
```


Job time : 182.319 secs

This Page Blank (usp10)

Db 1966 TCGCAGTCCACCACGGTGGTGTGATCGGCAGGACGGCGCAGGTTCCGGACGTGGAGATGCAG 1907

QY 21 TyrGlyAlaLeuCysLeu-AsnThrArgTyrGlyThrThrLeuAspGlu-GluLysAlaA 40

Db 1906 TTCGGCGCGCTGGCGTGCACCGTGGCTACGGCATGACCCCTGGACGAAGGAGCGC 1847

QY 40 rgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnA 60

Db 1846 GCATCTGTAGCAGCGCGCGCAGCGCGCGCTCGCCCGGCGCTGGCGCGCGAGCAGCAGC 1787

QY 60 rgLeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76

Db 1786 GCGTGGCGGACGGCGAGGAGGCGCGCGCTCTGGACGGAGGCGGAGAAG 1737

RESULT 2

US-08-933-750C-90

; Sequence 90, Application US/08933750C

; Patent No. 5932442

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Bandman, Olga

; APPLICANT: Shah, Purvi

; APPLICANT: Au-Young, Janice

; APPLICANT: Yue, Henry

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

; NUMBER OF SEQUENCES: 98

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/933,750C

; FILING DATE: September 23, 1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0356 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 90:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1078 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: SININOT01

; CLONE: 2184712

US-08-933-750C-90

Alignment Scores:

Pred. No.: 0.485 Length: 1078

Score: 77.50 Matches: 24

Percent Similarity: 50.77% Conservative: 9

Best Local Similarity: 36.92% Mismatches: 15

Query Match: 19.82% Indels: 17

DB: 2 Gaps: 4

US-10-029-020-14_COPY_2650_2725 (1-76) x US-08-933-750C-90 (1-1078)

QY 23 AlaLeuCysLeuAsnThr-----ArgTyrGlyThrThrLeuAspGluGluLys 38

Db 25 TCCCTGTGTGATGAGACATCATCCCTCCAGGAGCAAGCGGGAAGTCTGGAGGACGCTGAG 84

QY 39 AlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGln 58

Db 85 GGGCGG-----AGGCGGAGAGGCGGAGCTCGCGATGATGATGCTCGG----- 126

QY 59 GlnArgLeuArgGluGlyGluGlyLeuArgAla-----Tip 71

Db 127 CAGGCTCTCGGGAAGGGAAGAGGAGGAAAGGGCCCAACCCCTGAAGAAGCAATACAGAA 186

QY 72 ThrGluGlyGluLys 76

Db 187 ACTGAAGGAGACAGA 201

RESULT 3

US-09-234-613-90

; Sequence 90, Application US/09234613

; Patent No. 6132973

; GENERAL INFORMATION:

; APPLICANT: Lal, Preeti

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Bandman, Olga

; APPLICANT: Shah, Purvi

; APPLICANT: Au-Young, Janice

; APPLICANT: Yue, Henry

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES

; NUMBER OF SEQUENCES: 98

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/234,613

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/933,750

; FILING DATE: September 23, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0356 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 90:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1078 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: SININOT01

; CLONE: 2184712

US-09-234-613-90


```
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "PR-1/luc construct"
US-09-223-535-1

Alignment Scores:
Pred. No.: 42.4 Length: 14113
Score: 73.50 Matches: 18
Percent Similarity: 45.10% Conservative: 5
Best Local Similarity: 35.29% Mismatches: 15
Query Match: 18.80% Indels: 13
DB: 3 Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-223-535-1 (1-14113)
QY 38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArg--- 56
Db 2283 AAGACGGCGCCCTGGACCTCGCCCGGAGCGCGCCACGAGGCCGCGCGGACC 2342
QY 57 -----GluGlnGlnArgLeuArgGluGly 64
Db 2343 CGCGCCACGGCCCGGACGGCGGACGAGCGCCAGAGCGCCGAGAGGCC 2402
QY 65 GluGluGlyLeuArgAlaTrpThrGluGlyGlu 75
Db 2403 GAGCGGCGCGTGAGGCTTGACGCTAGGGCAG 2435

RESULT 8
US-09-023-655-969
; Sequence 969, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 969:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
```

```
; CLONE: g1504025
US-09-023-655-969

Alignment Scores:
Pred. No.: 39.6 Length: 4121
Score: 68.50 Matches: 14
Percent Similarity: 58.54% Conservative: 10
Best Local Similarity: 34.15% Mismatches: 10
Query Match: 17.52% Indels: 7
DB: 4 Gaps: 1

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-023-655-969 (1-4121)
QY 34 LeuAspGluGluLysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAla 53
Db 1211 CTTGAACACGAGAG-----CGCAGGAAGGAGATCAAGAGGCC 1249
QY 54 TrpAlaArgGluGlnGlnArgLeuArgGluGlyGluGlyLeuArgAlaTrpThrGlu 73
Db 1250 TGGCACCGTGGCCAGAGGAGCTGCAAGAGCGGAGTCCAACCTGCCCAAGGCCAAGCAG 1309
QY 74 Gly 74
Db 1310 GGT 1312

RESULT 9
US-09-252-991A-12184
; Sequence 12184, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12184
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12184

Alignment Scores:
Pred. No.: 6.41 Length: 876
Score: 68.00 Matches: 20
Percent Similarity: 44.83% Conservative: 6
Best Local Similarity: 34.48% Mismatches: 32
Query Match: 17.39% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-252-991A-12184 (1-876)
QY 7 LeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGlnTyrGlyAlaLeuCysLeu 26
Db 431 CTGGTCGGTCGGCTGGTGGTTCACCGACACCGCGCGCTTCGCCTCGACCCCTGGCGATC 490
QY 27 AsnThrArgTyrGlyThrLeuAspGluGluLysAlaArgValLeuGluLeuAlaArg 46
Db 491 CTCACCGCAGCGCGCTGCGCTGGTGGAGGCGCTGCGGATCGCCCGGAGGTGATCGCC 550
QY 47 GlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArgLeuArgGluGly 64
Db 551 AACCGGATCATCCGCAACGAAGTGGTCAAGCGCGCAGAGGTCCCGCGAGGC 604

RESULT 10
US-09-252-991A-2453
; Sequence 2453, Application US/09252991A
; Patent No. 6551795
```



```

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2535
; LENGTH: 2205
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2535

Alignment Scores:
Pred. No.: 20.8 Length: 2205
Score: 68.00 Matches: 24
Percent Similarity: 50.72% Conservative: 11
Best Local Similarity: 34.78% Mismatches: 15
Query Match: 17.39% Indels: 19
DB: 4 Gaps: 3

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-252-991A-2535 (1-2205)

QY 15 ThrAspIleGlnLeuGlnTyrGlyAla-----LeuCysLeuAsnThrArgTyrGlyThr 32
Db 1276 ACCCTGCTCAAGCTGCGCTACGGCAGTCCCGAGGCGCTGCTATTCAACC----- 1323

QY 33 ThrLeuAspGluGluLysAlaArg-----ValLeuGlu 43
Db 1324 -----GAGGAAGTGGCGCGGGAATGGCCCTGGTGGCTGGGAAGTGGCCCTGGAG 1374

QY 44 LeuAlaArgGln-ArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArgLeuArgG1 63
Db 1375 CTGGCCAGGGAGAGGCGCGCGGATCCTGGCGCAGGACTACGAAGTCACTGCGGAA 1434

QY 63 uGlyGluGluGlyLeuArgAlaTrp 71
Db 1435 ATGCTTCGAAGCTCCGGAATGG 1459

RESULT 14
US-09-252-991A-2384/c
; Sequence 2384, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2384
; LENGTH: 2388
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2384

Alignment Scores:
Pred. No.: 23 Length: 2388
Score: 68.00 Matches: 24
Percent Similarity: 50.72% Conservative: 11
Best Local Similarity: 34.78% Mismatches: 15
Query Match: 17.39% Indels: 19
DB: 4 Gaps: 3

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-252-991A-2384 (1-2388)

QY 15 ThrAspIleGlnLeuGlnTyrGlyAla-----LeuCysLeuAsnThrArgTyrGlyThr 32
Db 1095 ACCCTGCTCAAGCTGCGCTACGGCAGTCCCGAGGCGCTGCTATTCAACC----- 1048

QY 33 ThrLeuAspGluGluLysAlaArg-----ValLeuGlu 43
Db 1047 -----GAGGAAGTGGCGCGGGAATGGCCCTGGTGGCTGGGAAGTGGCCCTGGAG 997
```

```

QY 44 LeuAlaArgGln-ArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArgLeuArgG1 63
Db 996 CTGGCCAGGGAGAGGCGCGCGGATCCTGGCGCAGGACTACGAAGTCACTGCGGAA 937

QY 63 uGlyGluGluGlyLeuArgAlaTrp 71
Db 936 ATGCTTCGAAGCGTCCGGAATGG 912
```

```

RESULT 15
US-09-621-976-595
; Sequence 595, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 595
; LENGTH: 374
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 199..372
US-09-621-976-595
```

```

Alignment Scores:
Pred. No.: 2.93 Length: 374
Score: 67.00 Matches: 15
Percent Similarity: 68.00% Conservative: 2
Best Local Similarity: 60.00% Mismatches: 6
Query Match: 17.14% Indels: 2
DB: 4 Gaps: 1
```

US-10-029-020-14_COPY_2650_2725 (1-76) x US-09-621-976-595 (1-374)

```

QY 46 ArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArgLeuArgGluGlyGlu 65
Db 19 CGGGAGAGGCGAACTCGCGATGAGTGGTCTCGG-----CAGGCTCTTCGGGAAGGGGAA 72

QY 66 GluGlyLeuArgAla 70
Db 73 GAAGGAGAGAAAGGGCC 87
```

Search completed: August 14, 2004, 21:40:39
Job time : 34.4845 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: August 14, 2004, 01:13:26 ; Search time 891.313 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_2650_2725
Perfect score: 391
Sequence: 1 SQINTVLNGRTRRYTDIQLQ.....EQQLREGEGGLRAWTEGK 76
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1 13135 @runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOB=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	391	100.0	6246	29	AY413475	AY413475 Homo sapi
2	385	98.5	499	10	BF077525	BF077525 227453 MA
3	385	98.5	685	13	BU742156	BU742156 UI-E-EJ0-
4	384	98.2	5970	29	AY413476	AY413476 Pan trogl
5	371	94.9	881	13	BU384298	BU384298 603858267
6	371	94.9	976	13	BU365499	BU365499 603788050
7	368	94.1	621	14	CF182790	CF182790 UI-M-EY0-
8	365	93.4	913	13	BU387427	BU387427 603857864
9	364	93.1	693	13	BQ210132	BQ210132 UI-R-EP0-
10	358	91.6	514	9	AI176158	AI176158 EST219739
11	353	90.3	424	9	AV589580	AV589580 AV589580
12	344	88.0	748	9	AU169909	AU169909 AU169909
13	335	85.7	507	10	AW580211	AW580211 RC4-HT041
14	313	80.1	713	9	AL045769	AL045769 DKFZp434F
15	303	77.5	608	10	BB161536	BB161536 BB161536
16	303	77.5	949	29	CNS03WB0	AL263493 Tetraodon
17	303	77.5	3038	11	AK037897	AK037897 Mus muscu
18	299	76.5	769	13	BU454148	BU454148 603771046
19	299	76.5	863	13	BU382750	BU382750 603859889
20	299	76.5	946	13	BU132264	BU132264 603120954
21	296	75.7	311	9	AI184444	AI184444 qd40d07.x
22	296	75.7	359	9	AI222954	AI222954 qm30d03.x
23	296	75.7	378	9	AI200670	AI200670 qf94c12.x
24	296	75.7	380	9	AI887009	AI887009 wl95d07.x
25	296	75.7	382	9	AI199810	AI199810 qf95c06.x
26	296	75.7	382	9	AI274285	AI274285 ql61c09.x
27	296	75.7	382	12	BM717993	BM717993 UI-E-EJ0-
28	296	75.7	390	9	AI144798	AI144798 UI-R-BT0-
29	296	75.7	390	9	AW044155	AW044155 wy70e05.x
30	296	75.7	394	9	AI910459	AI910459 wi31c06.x
31	296	75.7	395	9	AI310370	AI310370 qo66h01.x
32	296	75.7	401	9	AI250756	AI250756 qi36b04.x
33	296	75.7	401	9	AI367099	AI367099 qk66e03.x
34	296	75.7	405	10	AW137000	AW137000 UI-H-BI1-
35	296	75.7	406	9	AI911042	AI911042 wd22g06.x
36	296	75.7	413	10	AW207016	AW207016 UI-H-BI1-
37	296	75.7	417	9	AI262318	AI262318 qk34c03.x
38	296	75.7	421	9	AI391741	AI391741 qy94cl2.x
39	296	75.7	430	9	AI338238	AI338238 qt41h06.x
40	296	75.7	440	10	AW137494	AW137494 UI-H-BI1-
41	296	75.7	447	9	AI827404	AI827404 wk94b10.x
42	296	75.7	450	9	AI288964	AI288964 ql81d06.x
43	296	75.7	453	9	AI278290	AI278290 ql78d04.x
44	296	75.7	454	9	AI079962	AI079962 oz34g12.x
45	296	75.7	458	9	AI017896	AI017896 ou35g05.x

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
6246 bp DNA linear GSS 12-DEC-2003
ACCESSION AY413475
VERSION AY413475.1 GI:39769437
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6246)

AUTHORS Clark,A.G., GJanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 6246)
AUTHORS Clark,A.G., GJanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
1..6246
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>6246
/locus_tag="HCM4903"
gene
ORIGIN
Alignment Scores:
Pred. No.: 9,11e-33 Length: 6246
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_2650_2725 (1-76) x AY413475 (1-6246)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
5884 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAG 5943
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
5944 TACGGGGCACTGTGCTTGAACACACGCTACGGGACAACGTTGGATGAGGAGGACGCGG 6003
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
6004 GTCCTGGAGCTGGCCCGCAGAGAGCCGTGCGCCACAGCGTGGCCCGCAGCAGAGA 6063
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
6064 CTGCGGGAAGGGAGGAGGAGGCGCTGCGGGCCTGGACAGAGGGGGAGAAG 6111
RESULT 2
BF077525 499 bp mRNA linear EST 25-APR-2001
LOCUS 227453 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION BF077525
ACCESSION BF077525.1 GI:10871355
VERSION EST.
KEYWORDS Bos taurus (cow)
SOURCE Bos taurus
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
1 (bases 1 to 499)
REFERENCE Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
AUTHORS Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
TITLE

JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
PUBMED 11282978
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 88 row: F column: 24
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1..499
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."
ORIGIN
Alignment Scores:
Pred. No.: 1.13e-33 Length: 499
Score: 385.00 Matches: 74
Percent Similarity: 100.00% Conservative: 2
Best Local Similarity: 97.37% Mismatches: 0
Query Match: 98.47% Indels: 0
DB: 10 Gaps: 0
US-10-029-020-14_COPY_2650_2725 (1-76) x BF077525 (1-499)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyThrAspIleGlnLeuGln 20
164 TCCAGATCAACACAGTACTTAACGGCAGGACTAGACGCTACACAGACATCCAGTCCAG 223
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
224 TATGGGGCTCTGTGCTTGAACACGCGCTACGGACCACTCTGGACGAGGAGAGCGCGG 283
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTyrAlaArgGluGlnGlnArg 60
284 GTGCTGGAGCTGGCCCGCAGAGAGCCGTGCGCCAGCGCTGGCCCGCTGGAGCAGCAGCA 343
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
344 CTTCGGACCGGGAGGAGGGCGCTGCGGGCCTGGACCGAGGGCGGAGAAG 391
RESULT 3
BU742156/c 685 bp mRNA linear EST 10-OCT-2002
LOCUS UI-E-EJ0-aim-m-24-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone
DEFINITION UI-E-EJ0-aim-m-24-0-UI 3', mRNA sequence.
ACCESSION BU742156
VERSION BU742156.1 GI:23688374
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 685)
REFERENCE Bonaldo,M.F., Lennon,G. and Soares,M.B.
AUTHORS Normalization and subtraction: two approaches to facilitate gene
TITLE discovery

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source

Location/Qualifiers
1. .685
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-aim-m-24-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ0"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ0 is a subtraced cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tags for this library are: fetal eyes,
AGATCAAGA; lens, CGATTAGCGA; eye anterior segment,
AATCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI).
TAG TISSUE=human fetal eyes
TAG_LIB=UI-E-EJ0
TAG_SEQ=AGATCAAGA"

ORIGIN

Alignment Scores:
Pred. No.: 1.79e-33 Length: 685
Score: 385.00 Matches: 75
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 1
Query Match: 98.47% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x BU742156 (1-685)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||||
Db 640 TCCAGATCAACACAGTACTTAATGCCAGGACTAGACGNTACAGACATCCAGCTCCAG 581
|||||
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
|||||
Db 580 TACGGGCACGTGCTTGAAACACACAGCTACGGGACACACCTTGGATGAGGAGGACCGG 521
|||||
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
|||||
Db 520 GTCCTGGAGCTGGCCCCGCGCAGAGAGCCGTGCGCCACAGCTGGGCCCGCGAGCAGAGAGA 461
|||||

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
|||||
Db 460 CTGCGGAAGGGAGGAGGCTGCGGGCCTGGACAGAGGGGAGAAAG 413
|||||

RESULT 4
AY413476
LOCUS
DEFINITION
AY413476 5970 bp DNA linear GSS 12-DEC-2003
Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY413476
VERSION
AY413476.1 GI:39769438
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 5970)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 5970)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .5970
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5970
/locus_tag="HCM4903"

gene

ORIGIN

Alignment Scores:
Pred..No.: 5.3e-32 Length: 5970
Score: 384.00 Matches: 75
Percent Similarity: 98.68% Conservative: 0
Best Local Similarity: 98.68% Mismatches: 1
Query Match: 98.21% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AY413476 (1-5970)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
|||||
Db 5608 TCCAGATCAACACAGTNCCTTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCAG 5667
|||||

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
|||||
Db 5668 TACGGGCACGTGCTTGAACACACACGCTACGGGACACACGTTGGATGAGGAGGACCGG 5727
|||||

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
|||||
Db 5728 GTCCTGGAGCTGGCCCCGCGCAGAGAGCCGTGCGCCAAAGCATGGCCCCGCGAGCAGAGAGA 5787
|||||

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
|||||
Db 5788 CTGCGGAANNNGAAGAGGCCTGCGGGCCTGGACAGAGGGGGAGAG 5835
|||||

RESULT 5
BU384298

LOCUS BU384298 881 bp mRNA linear EST 28-NOV-2002
DEFINITION 603858267F1 CSEQCHN75 Gallus gallus cDNA clone CHEST865f18 5', mRNA sequence.
ACCESSION BU384298
VERSION BU384298.1 GI:258922299
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 881)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..881
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="CHEST865f18"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHN75"
/note="Organ: trunks; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 9.93e-32 Length: 881
Score: 371.00 Matches: 71
Percent Similarity: 97.37% Conservative: 3
Best Local Similarity: 93.42% Mismatches: 2
Query Match: 94.88% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x BU384298 (1-881)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyTyrThrAspIleGlnLeuGln 20
Db 47 TCCAGATCAACACACAGTCTGGTGAAGGACTAGACGTTACACGGACATCCAGCTGCAG 106
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 107 TATGGTGGCGCTGTGTCTGAACACTCGCTACGGGACCACCTTTGGACGAGGAGAGGCCCGA 166
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 167 GTCCTGGAGCTGGCCCGGCAGCGCGCGTGGCCCAAGCTTGGCCCGGACACAGCAGAGA 226

QY 61 LeuArgGluGlyGluGlyLeuArgAlaTyrThrGluGlyGluLys 76
Db 227 TTGAGGGATGGGAGGAGGGGATTCGCTCGTGACAGAGGGGAGAAG 274
RESULT 6
BU365499 976 bp mRNA linear EST 28-NOV-2002
LOCUS 603788050F1 CSEQCHN72 Gallus gallus cDNA clone CHEST747cl3 5', mRNA sequence.
ACCESSION BU365499
VERSION BU365499.1 GI:25873500
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 976)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source
1..976
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST747cl3"
/sex="Female"
/tissue_type="cerebrum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN72"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 1.15e-31 Length: 976
Score: 371.00 Matches: 71
Percent Similarity: 97.37% Conservative: 3
Best Local Similarity: 93.42% Mismatches: 2
Query Match: 94.88% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x BU365499 (1-976)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgGlyTyrThrAspIleGlnLeuGln 20

Db	173	TCCACATCAACACAGTGCTGGGTGAAGGACTAGACGTTACACGGACATCCAGCTGCAG	232
QY	21	TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg	40
Db	233	TACGGTCGGCTGTGTCTGAACACTCGCTACGGGACCACCTTTGGACGAGGAGGCCCGA	292
QY	41	ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg	60
Db	293	GTCTGGAGCTGGCCCGCAGCGCGCTGGCCCAAGCTTGGCCCGGGAACAGCAGAGA	352
QY	61	LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys	76
Db	353	TTGAGGATGGGAGGAGGGGATTCGCTCGTGACAGAGGGGAGAAG	400
RESULT 7			
CF182790			
LOCUS			
DEFINITION	CF182790	621 bp mRNA linear	EST 29-JUL-2003
ACCESSION	UI-M-EY0-bwz-p-16-0-UI.r1	NIH_BMAP_EY0	Mus musculus cDNA clone
VERSION	IMAGE: 5705343	5', mRNA sequence.	
KEYWORDS	CF182790		
SOURCE	CF182790.1	GI:33314674	
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 621)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Dr. James Lin, Univeristy of Iowa		
	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
	DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
	Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	http://image.llnl.gov		
	This clone was contributed by the Brain Molecular Anatomy Project (BMAP)		
Seq primer: pYX-5.			
FEATURES	Location/Qualifiers		
source	1. .621		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6"		
	/db_xref="taxon:10090"		
	/clone="IMAGE: 5705343"		
	/tissue_type="whole brain"		
	/dev_stage="embryo 15.5 dpc"		
	/lab_host="DH10B (r1 phage resistant)"		
	/clone_lib="NIH_BMAP_EY0"		
	/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is GTGGCTGGAA. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.31e-31	Length:	621
Score:	368.00	Matches:	71

Percent Similarity:	97.37%	Conservative:	3
Best Local Similarity:	93.42%	Mismatches:	2
Query Match:	94.12%	Indels:	0
DB:	14	Gaps:	0
US-10-029-020-14_COPY_2650_2725 (1-76) x CF182790 (1-621)			
QY	1	SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln	20
Db	377	TCCAGATCAACACCGTGTCTAGTGGCAGGACTAGACGCTACACTGACATCCAGCTGCAG	436
QY	21	TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg	40
Db	437	TACAGGGCGCTGTGCCTGAACACCCGCTACGGGACACAGTGGATGAGGAAAGGTGCGG	496
QY	41	ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg	60
Db	497	GTCTGGAGCTGGCCAGGCAGAGAGCTGTGCGCCAGCGTGGCCCGGAGCAGCAGAGA	556
QY	61	LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys	76
Db	557	CTGCGGGAAGGGAGAGGGCCTCGGGCCCTGGACAGATGGGGAAG	604

RESULT 8			
BU387427			
LOCUS	BU387427	913 bp mRNA linear	EST 28-NOV-2002
DEFINITION	603857864F1 CSEQCHN75	Gallus gallus cDNA clone	ChEST864p16 5', mRNA sequence.
ACCESSION	BU387427		
VERSION	BU387427.1	GI:25895415	
KEYWORDS	EST.		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
AUTHORS	1 (bases 1 to 913)		
TITLE	Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.		
JOURNAL	A Comprehensive Collection of Chicken cDNAs		
MEDLINE	Curr. Biol. 12 (22), 1965-1969 (2002)		
PUBMED	22335534		
COMMENT	12445392		
	Contact: Simon Hubbard		
	Department of Biomolecular Sciences		
	University of Manchester Institute of Science and Technology (UMIST)		
	PO Box 88, Manchester, M60 1QD, UK		
	Tel: 01612008930		
	Fax: 01612360409		
	Email: Simon.Hubbard@umist.ac.uk .		
FEATURES	Location/Qualifiers		
source	1. .913		
	/organism="Gallus gallus"		
	/mol_type="mRNA"		
	/strain="White Leghorn, Hisex"		
	/db_xref="taxon:9031"		
	/clone="ChEST864p16"		
	/dev_stage="36"		
	/lab_host="DH10B"		
	/clone_lib="CSEQCHN75"		
	/note="Organ: trunks; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was bluntend, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6		


```
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
Location/Qualifiers
1..514
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="ATCC (inhost):2035740"
/db_xref="taxon:10118"
/clone="ROVEL55"
/clone_lib="Normalized rat ovary, Bento Soares"
/note="Organ: ovary; Vector: pT7T3Pac; Site_1: EcoRI;
Site_2: NotI"

ORIGIN
Alignment Scores:
Pred. No.: 1.35e-30 Length: 514
Score: 358.00 Matches: 69
Percent Similarity: 94.74% Conservative: 3
Best Local Similarity: 90.79% Mismatches: 4
Query Match: 91.56% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x A1176158 (1-514)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 443 TCCCAAATCAACCCCATGCTCGGTGGCAGGACTAGACGCTACACTGACATCCAGCTGCAG 384
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 383 TACAGGGCACTGTGCTGAACACCCCGCTATGGACAACCGTGGACGAGGAAAGGTACGG 324
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 323 GTGCTGGAGCTGGCCAGGCAGAGGGCTGTACCCAGGCCTGGGCCCGGAGCAGCAGAGA 264
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 263 CTGCGGGAAGGGGAAGAGGGCCTGCGGGCTGGACAGACGGGGAGAAG 216

RESULT 11
AV589580/c
LOCUS AV589580 Bos taurus brain fetus Bos taurus cDNA clone E1BR007C10
DEFINITION AV589580 Bos taurus brain fetus Bos taurus cDNA clone E1BR007C10
3', mRNA sequence.
ACCESSION AV589580
VERSION AV589580.1 GI:9700573
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 424)
Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
and Sugimoto,Y.
Establishment of a high throughput EST sequencing system using
poly(A) tail-removed cDNA libraries and determination of 36,000
bovine ESTs
Nucleic Acids Res. 29 (22), E108 (2001)
MEDLINE 21570554
PUBMED 11713328
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.
Location/Qualifiers

FEATURES
source
1..424
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR007C10"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus brain fetus"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

ORIGIN
Alignment Scores:
Pred. No.: 3.78e-30 Length: 424
Score: 353.00 Matches: 68
Percent Similarity: 94.67% Conservative: 3
Best Local Similarity: 90.67% Mismatches: 4
Query Match: 90.28% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AV589580 (1-424)
QY 2 GlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGlnTyr 21
Db 422 CAGATCAACCCGATGCTTAACGGCAGGACTAGACGCTACNAGACATCCAGCTCCAGTAT 363
QY 22 GlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArgVal 41
Db 362 GGGGCTCTGTGCTTGAACACCGCTACGGACCACTCTGNACGAGAAGAAGCGGGGTG 303
QY 42 LeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArgLeu 61
Db 302 CTGGAGCTGGCCGCGCAGAGAGCCGTGCGCAGGCCTGGGCCCTGTGAGCAGCGACTT 243
QY 62 ArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 242 CGGACCGGNAGGAGGGCCTGCGGGCTGGACCGAGGCGGAGAAG 198

RESULT 12
AUI69909/c
LOCUS AUI69909 Ol-br-ad cDNA Oryzias latipes cDNA clone br5373, mRNA
DEFINITION AUI69909 Ol-br-ad cDNA Oryzias latipes cDNA clone br5373, mRNA
sequence.
ACCESSION AUI69909
VERSION AUI69909.1 GI:12591978
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 748)
Mita,K., Ishikawa,Y. and Yamauchi,M.
Establishment of cDNA database of medaka, Oryzias latipes
Unpublished (2001)
Contact: Mita K
Genome Research Group
National Institute of Agrobiological Sciences
Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan
Email: kmita@nias.affrc.go.jp
method:uni-directional sequence direction:sequenced from T3 primer
(5' -> 3').

FEATURES
source
1..748
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="HNI"
/db_xref="taxon:8090"
/clone="br5373"
/sex="female/male mixed"
/tissue_type="brain"
```

```
source
1..424
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="E1BR007C10"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH10B"
/clone_lib="Bos taurus brain fetus"
/note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
was deleted from a NotI site"

ORIGIN
```

```
Alignment Scores:
Pred. No.: 3.78e-30 Length: 424
Score: 353.00 Matches: 68
Percent Similarity: 94.67% Conservative: 3
Best Local Similarity: 90.67% Mismatches: 4
Query Match: 90.28% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AV589580 (1-424)
QY 2 GlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGlnTyr 21
Db 422 CAGATCAACCCGATGCTTAACGGCAGGACTAGACGCTACNAGACATCCAGCTCCAGTAT 363
QY 22 GlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArgVal 41
Db 362 GGGGCTCTGTGCTTGAACACCGCTACGGACCACTCTGNACGAGAAGAAGCGGGGTG 303
QY 42 LeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArgLeu 61
Db 302 CTGGAGCTGGCCGCGCAGAGAGCCGTGCGCAGGCCTGGGCCCTGTGAGCAGCGACTT 243
QY 62 ArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 242 CGGACCGGNAGGAGGGCCTGCGGGCTGGACCGAGGCGGAGAAG 198
```

```
RESULT 12
AUI69909/c
LOCUS AUI69909 Ol-br-ad cDNA Oryzias latipes cDNA clone br5373, mRNA
DEFINITION AUI69909 Ol-br-ad cDNA Oryzias latipes cDNA clone br5373, mRNA
sequence.
ACCESSION AUI69909
VERSION AUI69909.1 GI:12591978
KEYWORDS EST.
SOURCE Oryzias latipes (Japanese medaka)
ORGANISM Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 748)
Mita,K., Ishikawa,Y. and Yamauchi,M.
Establishment of cDNA database of medaka, Oryzias latipes
Unpublished (2001)
Contact: Mita K
Genome Research Group
National Institute of Agrobiological Sciences
Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan
Email: kmita@nias.affrc.go.jp
method:uni-directional sequence direction:sequenced from T3 primer
(5' -> 3').

FEATURES
source
1..748
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="HNI"
/db_xref="taxon:8090"
/clone="br5373"
/sex="female/male mixed"
/tissue_type="brain"
```

```
FEATURES
source
1..748
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="HNI"
/db_xref="taxon:8090"
/clone="br5373"
/sex="female/male mixed"
/tissue_type="brain"
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 128.916 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_2650_2725
Perfect score: 391
Sequence: 1 SQINTVLNGRTRRYTDIQLQ.....EQRLREGEGGLRAWTEGK 76

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_1868@runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:
1: geneseqn1980s:
2: geneseqn1990s:
3: geneseqn2000s:
4: geneseqn2001as:
5: geneseqn2001bs:
6: geneseqn2002s:
7: geneseqn2003as:
8: geneseqn2003bs:
9: geneseqn2003cs:
10: geneseqn2004s:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	Score	Match	Length	DB	ID	Description
No.							
1	391	100.0	8354	6	ABS52100	Abs52100	Human TEN
2	391	100.0	8438	6	ABN85378	Abn85378	Human NOV
3	391	100.0	8645	6	ABS78652	Abs78652	Human CDN
4	311	79.5	3312	4	AAK52812	Aak52812	Human pol
5	311	79.5	13202	4	AAK51828	Aak51828	Human pol
6	308	78.8	12879	6	ABK92230	Abk92230	Prostate
7	296	75.7	2157	4	AAH14096	Aah14096	Human CDN
8	296	75.7	3270	4	AAH14671	Aah14671	Human CDN

9	296	75.7	3614	4	AAH14183	Aah14183 Human cdn
10	296	75.7	8473	6	ABQ82345	Abq82345 Human NOV
11	296	75.7	8487	6	ABQ82346	Abq82346 Human NOV
12	296	75.7	8645	6	ABQ82344	Abq82344 Human NOV
13	296	75.7	8675	6	ABQ82343	Abq82343 Human NOV
14	246.5	63.0	9058	7	ACC72051	Acc72051 BCU0205A
15	246.5	63.0	9695	7	ACC72052	Acc72052 BCU0205B
16	246.5	63.0	9729	5	AAS14089	Aas14089 Human FCT
17	246.5	63.0	9729	9	ADB32028	Adb32028 Human FCT
18	246.5	63.0	9826	5	AAS14085	Aas14085 Human FCT
19	246.5	63.0	9826	9	ADB32023	Adb32023 Human FCT
20	135.5	34.7	540	4	AAH10637	Aah10637 Human cdn
21	124	31.7	3910	4	ABL04849	Ab104849 Drosophil
22	124	31.7	6318	4	ABL04848	Ab104848 Drosophil
23	117.5	30.1	464	8	ACH40709	Ach40709 Human foe
24	103	26.3	10242	4	ABL29075	Ab129075 Drosophil
25	103	26.3	17131	4	ABL29074	Ab129074 Drosophil
26	90	23.0	65	6	ABN55563	Abn55563 Mouse spl
27	77.5	19.8	372	3	AAC08271	Aac08271 Human sec
28	77.5	19.8	458	3	AAC01764	Aac01764 Human sec
29	77.5	19.8	2015	10	ADE77097	Ade77097 Human cdn
30	75.5	19.3	992	6	ABK34678	Abk34678 Human cdn
31	75	19.2	6457	6	ABN96892	Abn96892 Gene #339
32	73.5	18.8	285	6	ABL69281	Ab169281 Prostate
33	73.5	18.8	445	6	ABT04001	Abt04001 Human ova
34	73.5	18.8	791	6	ABK35440	Abk35440 Human cdn
35	73.5	18.8	14113	3	AAA39651	Aaa39651 Fusion co
36	73.5	18.8	14113	3	AAA90096	Aaa90096 PR-1 prom
37	73.5	18.8	14113	4	AAH25853	Aah25853 PR-1 prom
38	73	18.7	60	6	ABN43073	Abn43073 Human spl
39	72	18.4	15231	6	ABK63598	Abk63598 Rat seque
40	72	18.4	15231	7	ABT42078	Abt42078 Toxicity
41	72	18.4	15231	9	ADB59205	Adb59205 Toxicity-
42	72	18.4	15231	9	ADB53026	Adb53026 Primary r
43	70.5	18.0	300	2	AAZ13121	Aaz13121 Human gen
44	70.5	18.0	761	6	ABQ60799	Abq60799 Human mrn
45	70.5	18.0	775	2	AAZ15180	Aaz15180 Human gen

ALIGNMENTS

RESULT 1

ABS52100

ID ABS52100 standard; DNA; 8354 BP.

XX

AC ABS52100;

XX

DT 05-NOV-2002 (first entry)

XX

DE Human TEN-M4-like gene.

XX

KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type Ia membrane sushi-containing domain; butyrophilin;
KW single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers

variation replace(117,G)

FT /*tag= a

FT /standard name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
XX 25-JUL-2002.
XX
XX 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
PR 20-DEC-2000; 2000US-0257314P.
PR 02-MAY-2001; 2001US-0288153P.
PR 29-MAY-2001; 2001US-0294075P.
PR 24-JUL-2001; 2001US-0307506P.
PR 10-AUG-2001; 2001US-0311590P.
PR 10-AUG-2001; 2001US-0311613P.
PR 29-AUG-2001; 2001US-0315617P.
PR 14-SEP-2001; 2001US-0322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
DR P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.
XX
XX Claim 8; Page 50-52; 318pp; English.
XX
XX The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders. The
CC present nucleic acid sequence represents a NOVX gene. This sequence
CC encodes a NOVX protein of the invention
XX
SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.53e-39 Length: 8354
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABS52100 (1-8354)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db TCCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGACATCCAGTCCAG 8041

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 8042 TACGGGGCACTGTGCTTGAACACACACGCTACGGGACACGTTGGATGAGGAGGACCG 8101

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 8102 GTCTGGAGCTGGCCCCGGCAGAGAGCCGTGCGCCAAAGCGTGGCGCCGCGAGCAGAGA 8161

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8162 CTGCGGGAAGGGGAGGAGGAGGCTGCGGGCCTGGACAGAGGGGAGAGA 8209

RESULT 2
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
XX
AC ABN85378;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human NOV1, TEN-M4 like protein, coding sequence.
XX
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 4..8395
FT /*tag= a
FT /trans_except= (pos: 1138..1147,aa:Met)
FT /product= "NOV1 protein"
XX
WO200255704-A2.
PN
XX 18-JUL-2002.
XX
XX 09-JAN-2002; 2002WO-US0000554.
XX
XX 09-JAN-2001; 2001US-0260417P.
PR 10-JAN-2001; 2001US-0260831P.
PR 28-FEB-2001; 2001US-0272338P.
PR 09-MAR-2001; 2001US-0274876P.
PR 18-APR-2001; 2001US-0284704P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Padigar M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
PI Gunther E, Stone DJ;
XX
XX WPI; 2002-590674/63.
DR P-PSDB; ABB98401.
XX
XX NOVX polypeptides and encoding polynucleotides, useful for preventing or
PT treating NOVX-associated disorders e.g. cancer, inflammation, or
PT Alzheimer's disease, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 9; Page 8-9; 358pp; English.
PS The present sequence is a coding sequence for a NOV protein. The NOV
XX
CC

CC proteins and coding sequences are useful for treating or preventing NOV-
CC associated disorders or in the manufacture of a medicament for treating
CC the disorders, such as cancer, heart disease, inflammation, autoimmune
CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
CC and other wasting disorders associated with chronic diseases. NOV1 is a
CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX

SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.59e-39 Length: 8438
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABN85378 (1-8438)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 8033 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCAG 8092
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8093 TACGGGGCACTGTGCTTGAACACACACGCTACGGGACAACGTTGGATGAGGAGGCACGG 8152
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 8153 GTCCTGGAGCTGGCCCGCAGAGAGCCGTGCGCAAGCGTGGGCCCGCGAGCAGAGAGA 8212
QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8213 CTGCGGGAAGGGAGGAAGGCCTGCGGGCCTGGACAGAGGGGAGAG 8260

RESULT 3
ABS78652

ID ABS78652 standard; cDNA; 8645 BP.

XX ABS78652;

XX 16-DEC-2002 (first entry)

XX Human cDNA encoding CGDD10, INCYTE 7488573CB1.

KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.

XX Homo sapiens.

XX WO200272830-A2.

XX 19-SEP-2002.

XX 08-FEB-2002; 2002WO-US003715.

XX 09-FEB-2001; 2001US-0268111P.

PR 23-FEB-2001; 2001US-0271175P.

PR 08-MAR-2001; 2001US-0274503P.

PR 09-MAR-2001; 2001US-0274552P.

XX

(INCY-) INCYTE GENOMICS INC.

PA Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
XX Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
PI WPI; 2002-723356/78.
XX P-PSDB; ABG97359.

PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.

XX Claim 5; Page 175-178; 181pp; English.

CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
XX a CGDD protein

SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 4.73e-39 Length: 8645
Score: 391.00 Matches: 76
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABS78652 (1-8645)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 8031 TCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCAG 8090
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8091 TACGGGGCACTGTGCTTGAACACACACGCTACGGGACAACGTTGGATGAGGAGGCACGG 8150
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 8151 GTCCTGGAGCTGGCCCGCAGAGAGCCGTGCGCAAGCGTGGGCCCGCGAGCAGAGAGA 8210
QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8211 CTGCGGGAAGGGAGGAAGGCCTGCGGGCCTGGACAGAGGGGAGAG 8258

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 3017 TCCAGATGACTTCTGTGTGAATGGAGGACTAGACGGTTTGACAGATATTCCAGCTCCAG 3076
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrLeuAspGluGluLysAlaArg 40
Db 3077 CATGGAGCCCTGTGCTTCAACATCCGGTATGGGACAACTGTCGAAGAGGAAAGAAATCAC 3136
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 3137 GTGTGGAGATTGCCAGACAGCGCGCAGTGGCCCGCTGGAAGGAAAGGAGG 3196
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 3197 CTGCAAGAGGGGGAAGAGGGGATTAGGGCATGGACAGAGAGGAGGAAAG 3244
RESULT 5
AAK51828
ID AAK51828 standard; cDNA; 13202 BP.
XX
AC AAK51828;
DT
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 373.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM78695.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 1; Page 1414-1426; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

RESULT 4
AAK52812
ID AAK52812 standard; cDNA; 3312 BP.
XX
AC AAK52812;
DT
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 2341.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR P-PSDB; AAM79679.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 1; Page 4651-4652; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 3312 BP; 928 A; 750 C; 739 G; 895 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2.23e-29 Length: 3312
Score: 311.00 Matches: 57
Percent Similarity: 89.47% Conservative: 11
Best Local Similarity: 75.00% Mismatches: 8
Query Match: 79.54% Indels: 0
DB: 4 Gaps: 0
US-10-029-020-14_COPY_2650_2725 (1-76) x AAK52812 (1-3312)

CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.26e-28 Length: 13202
Score: 311.00 Matches: 57
Percent Similarity: 89.47% Conservative: 11
Best Local Similarity: 75.00% Mismatches: 8
Query Match: 79.54% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AAK51828 (1-13202)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 8207 TCCAGATGACTTCTGTGTTGAATGGGAGGACTAGACGGTTTGCAGATATTCAGCTCCAG 8266
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8267 CATGGAGCCCTGTGTTCAACATCCGGTATGGGACAACTGTCGAAGAGGAAAGAAATCAC 8326
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 8327 GTGTTGGAGATTGCCAGACAGCGCGAGTGGCCAGGCCTGGACTAAGGAACAAAGAAGG 8386
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8387 CTGCAAGAGGGGGAAGAGGGGATTAGGGCATGGACAGAGGGGGAAG 8434

RESULT 6
ABK92230
ID ABK92230 standard; DNA; 12879 BP.
XX
AC ABK92230;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated DNA sequence #116.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene; ds.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX

PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
(EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
DR WPI; 2002-471335/50.
DR P-PSDB; ABG61913.
XX

Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a

PT prostate tissue.
XX
PS Claim 22; Page 394-397; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences

XX
SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.96e-28 Length: 12879
Score: 308.00 Matches: 56
Percent Similarity: 89.47% Conservative: 12
Best Local Similarity: 73.68% Mismatches: 8
Query Match: 78.77% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABK92230 (1-12879)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db 7880 TCCAGATGACTTCTGTGTTGAATGGGAGGACTAGACGGTTTGCAGATATTCAGCTCCAG 7939
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 7940 CATGGAGCCCTGTGTTCAACATCCGGTATGGGACAACTGTGCAAGAGGAAAGAAATCAC 7999
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 8000 GTGTTGGAGATTGCCAGACAGCGCGCAGTGGCCAGGCCTGGACTAAGGAACAAAGAAGG 8059
QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 8060 CTGCAAGAGGGGGAAGAGGGGATTAGGGCATGGACAGAGGGGGAAG 8107

RESULT 7
AAH14096
ID AAH14096 standard; cDNA; 2157 BP.
XX
AC AAH14096;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:11265.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.

PR	09-JUN-2000; 2000JP-00241899.	DT	26-JUN-2001 (first entry)
XX		XX	Human cDNA sequence SEQ ID NO:12354.
PA	(HELI-) HELIX RES INST.	DE	
XX		XX	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	XX	Homo sapiens.
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	XX	
XX		XX	EP1074617-A2.
DR	WPI; 2001-318749/34.	PN	
XX		XX	07-FEB-2001.
XX		XX	28-JUL-2000; 2000EP-00116126.
PT	Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.	XX	29-JUL-1999; 99JP-00248036.
XX		PR	27-AUG-1999; 99JP-00300253.
PS	Claim 8; SEQ ID NO 11265; 2537pp + Sequence Listing; English.	PR	11-JAN-2000; 2000JP-00118776.
XX		PR	02-MAY-2000; 2000JP-00183767.
XX		PR	09-JUN-2000; 2000JP-00241899.
XX		XX	(HELI-) HELIX RES INST.
XX		XX	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX		PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX		XX	WPI; 2001-318749/34.
XX		DR	
XX		XX	Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX		XX	Claim 8; SEQ ID NO 12354; 2537pp + Sequence Listing; English.
XX		XX	The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX		XX	Sequence 2157 BP; 544 A; 560 C; 591 G; 462 T; 0 U; 0 Other;
SQ		SQ	Sequence 3270 BP; 867 A; 821 C; 852 G; 730 T; 0 U; 0 Other;
			Alignment Scores:
			Pred. No.: 1.06e-27 Length: 2157
			Score: 296.00 Matches: 54
			Percent Similarity: 86.84% Conservative: 12
			Best Local Similarity: 71.05% Mismatches: 10
			Query Match: 75.70% Indels: 0
			DB: 4 Gaps: 0
			US-10-029-020-14_COPY_2650_2725 (1-76) x AAH14096 (1-2157)
QY	1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20	QY	1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db	1550 TCGCAGTCCACCGGTGTGAACGGCAGGACCGCGCGCTTCGGGACGTGGAGATGCAG 1609	Db	2580 TCGCAGTCCACCGGTGTGAACGGCAGGACCGCGCGCTTCGGGACGTGGAGATGCAG 2639
QY	21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40	QY	1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db	1610 TTCGGCGCGTGGCGCTGCAGTGCCTACGGCATGACCTGGACGAGGAGGCGCGC 1669	Db	2580 TCGCAGTCCACCGGTGTGAACGGCAGGACCGCGCGCTTCGGGACGTGGAGATGCAG 2639
QY	41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnArg 60	QY	1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db	1670 ATCCTGGACGAGCGCGCGCTCGCCCGGCTTCGGGCGCGGAGCAGCGC 1729	Db	2580 TCGCAGTCCACCGGTGTGAACGGCAGGACCGCGCGCTTCGGGACGTGGAGATGCAG 2639
QY	61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76	QY	1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgTyrThrAspIleGlnLeuGln 20
Db	1730 GTGGCGACGGCAGGAGGGCGCGCTCTGTGACGGAGGGCGAGAG 1777	Db	2580 TCGCAGTCCACCGGTGTGAACGGCAGGACCGCGCGCTTCGGGACGTGGAGATGCAG 2639
			RESULT 8
			AAH14671
			ID AAH14671 standard; cDNA; 3270 BP.
			XX
			AC AAH14671;
			XX

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db ::
2640 TTCGGCGGCTGGCGCTGCAGTGGCTACGGCATGACCTGGACGAGAGGCGCGC 2699

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db ::
2700 ATCCTGGAGCAGCGCGCAGCGCGCTCGCCGCGCTGGCGCGCGAGCAGCAGCGC 2759

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db ::
2760 GTGCGCGACGCGGAGGCGCGCGCTCTGGACGAGGGCGGAGAAG 2807

RESULT 9
AAH14183
ID AAH14183 standard; cDNA; 3614 BP.
XX
AC AAH14183;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:11430.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
DR WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 11430; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the

CC present invention
XX
SQ Sequence 3614 BP; 967 A; 898 C; 928 G; 821 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.03e-27 Length: 3614
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x AAH14183 (1-3614)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db ::
2924 TCGCAGTCCACCAACGCGGTGAACGGCAGGACGCGCAGGTTTCGCGGACGTGGAGATGCAG 2983

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db ::
2984 TTCGGCGCGCTGGCGCTGCACGTGCGCTACGCGATGACCTGGACGAGAGAGGCGCGC 3043

QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db ::
3044 ATCCTGGAGCAGCGCGCGCGCTCGCCCGGCGCTGGCGCGCGGAGGCGGAG 3103

QY 61 LeuArgGluGlyGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db ::
3104 GTGCGCGACGCGGAGGAGGGCGCGCTCTGGACGAGGGCGGAGAG 3151

RESULT 10

ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX
AC ABQ82345;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15c encoding cDNA SEQ ID NO:39.

Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; Pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; ss.

Homo sapiens.

Key Location/Qualifiers
CDS 258..8144
/*tag= a
/product= "NOV15c"

WO200262999-A2.

15-AUG-2002.

31-DEC-2001; 2001WO-US049976.

29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.

PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53589.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 119-121; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where x is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15C, which is
CC located on chromosome 4
XX
SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 5.92e-27 Length: 8473
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABQ82345 (1-8473)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7782 TCGAGTCCACACACCGTGTGTGAACCGCAGACGCGAGTTCCGCGACGTGGAGATGCAG 7841

QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 7842 TTCGGCGCGTGGCGCTGCACGTGCGCTACGCGCATGACCCCTGGACGAGGAGAGCGCGC 7901

QY 41 ValLeuGluLeuAlaArgGlnAlaValArgGlnAlaTrpAlaArgGluGlnArg 60
Db 7902 ATCCTGGAGCAGCGCGCGCAGCGCGCTCGCCCGCGCCTGGCGCGCAGCAGCAGCGC 7961

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 7962 GTGCGGACGGCGAGGAGGGCGCGCCTCTGACGAGGAGGCGAGAAG 8009
RESULT 11
ABQ82346
ID ABQ82346 standard; cDNA; 8487 BP.
XX
AC ABQ82346;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15d encoding cDNA SEQ ID NO:41.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 299..8140
FT /*tag= a
FT /product= "NOV15d"
XX
PN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53589.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 123-125; 444pp; English.

SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 6.07e-27 Length: 8645
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABQ82344 (1-8645)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7954 TCGCAGTCCACCACGGTGGTGAACGCGAGCAGCGCAGGTTTCGCGGACGTGGAGATGCAG 8013
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8014 TTCGGCGCGCTGGCGCTGCACGTGCGCTACGGCATGACCCCTGGAYGAGGAGAGCGCGC 8073
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 8074 ATCCTGGAGCAGCGCGCGCAGCGCGCTGCGCGCGCGCTGGCGCGCGAGCAGCAGCGC 8133
QY 61 LeuArgGluGluGluGluGluLeuArgAlaTrpThrGluGlyGluLys 76
Db 8134 GTGCGCGACGCGAGGAGCGCGCGCGCTCTGGACGGAGCGCGAGAAG 8181

RESULT 13

ABQ82343
ID ABQ82343 standard; cDNA; 8675 BP.

XX ABQ82343;

DT 17-DEC-2002 (first entry)

XX Human NOV15a encoding cDNA SEQ ID NO:35.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 151..8328
FT /*tag= a
FT /product= "NOV15a"
FT /transl_except= (pos:1249..1251,aa:Ser)

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.

PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX (CURA-) CURAGEN CORP.
XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
XX Gunther E, Smithson G, Millet I, Macdougall JR;
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53586.
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX Claim 8; Page 110-112; 444pp; English.
PS

XX The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastric and duodenal disorders),
CC disorders (e.g. ulcerative colitis, or allergic reactions, autoimmune haemolytic
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4

XX SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 6.09e-27 Length: 8675
Score: 296.00 Matches: 54
Percent Similarity: 86.84% Conservative: 12
Best Local Similarity: 71.05% Mismatches: 10
Query Match: 75.70% Indels: 0
DB: Gaps: 0

US-10-029-020-14_COPY_2650_2725 (1-76) x ABQ82343 (1-8675)

QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7966 TCGCAGTCCACCACGGTGGTGAACGCGAGCAGCGCAGGTTTCGCGGACGTGGAGATGCAG 8025
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGlyThrThrLeuAspGluGluLysAlaArg 40
Db 8026 TTCGGCGCGCTGGCGCTGCACGTGCGCTACGGCATGACCCCTGGACGAGAGGCGCGC 8085
QY 41 ValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGluGlnGlnArg 60
Db 8086 ATCCTGGAGCAGCGCGCGCAGCGCGCTGCGCGCGCTGGCGCGCGAGCAGCAGCGC 8145

QY 61 LeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
ACC72051
Db 8146 GTGCGGACGGCGAGGAGGGCGCGCTCTGGACGGAGGGGAGAAG 8193
RESULT 14
ID ACC72051 standard; DNA; 9058 BP.
XX
AC ACC72051;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205A gene #SEQ ID 79.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
PN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58317.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.31e-20 Length: 9058
Score: 246.50 Matches: 49
Percent Similarity: 75.95% Conservative: 11
Best Local Similarity: 62.03% Mismatches: 16
Query Match: 63.04% Indels: 3
DB: 7 Gaps: 1
US-10-029-020-14_COPY_2650_2725 (1-76) x ACC72051 (1-9058)
QY 1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db 7328 TCCAGCCCCACGCTGCTGTCAACGGCAGGACTCGAAGTTTCACGAACATTGAGTCCAG 7387
QY 21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37

Db 7388 TACTCCACGCTGCTGCTCAGCATCGCTATGGCCTACCCCCGACACCCCTGGACGAAG 7447
QY 38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlu 57
Db 7448 AAGGCCCGCGTCTCTGGACCGAGAGACAGAGGGCCCTGGGCACGGCCTGGCCCAAGGAG 7507
QY 58 GlnGlnArgLeuArgGluGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db 7508 CAGCAGAAAGCCAGGACGGAGAGAGGGGAGCGCCCTGTGGACTGAGGGCGAGAAG 7564
RESULT 15
ACC72052
ID ACC72052 standard; DNA; 9695 BP.
XX
AC ACC72052;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205B gene #SEQ ID 81.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
PN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58318.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.43e-20 Length: 9695
Score: 246.50 Matches: 49
Percent Similarity: 75.95% Conservative: 11
Best Local Similarity: 62.03% Mismatches: 16
Query Match: 63.04% Indels: 3
DB: 7 Gaps: 1

```
US-10-029-020-14_COPY_2650_2725 (1-76) x ACC72052 (1-9695)
QY      1 SerGlnIleAsnThrValLeuAsnGlyArgThrArgArgTyrThrAspIleGlnLeuGln 20
Db      7965 TCCAGCCCAAGCTGCTGCTCAACGGCAGGACTCGAAGGTTACGAACATTGAGTTCCAG 8024
QY      21 TyrGlyAlaLeuCysLeuAsnThrArgTyrGly-----ThrThrLeuAspGluGlu 37
Db      8025 TACTCCAGCTGCTGCTCAGCATCCGCTATGGCCTCACCCCGCACACCTGGACGAGAG 8084
QY      38 LysAlaArgValLeuGluLeuAlaArgGlnArgAlaValArgGlnAlaTrpAlaArgGlu 57
Db      8085 AAGCCCGCGTCTGGACCAAGCGAGACAGAGGGCCCTGGGCACGGCCTGGGCCAAGGAG 8144
QY      58 GlnGlnArgLeuArgGluGlyGluGluGlyLeuArgAlaTrpThrGluGlyGluLys 76
Db      8145 CAGCAGAAAGCCAGGACGGGAGAGAGGGGAGCCGCCCTGTGGACTGAGGGCGGAGAAG 8201
```

Search completed: August 14, 2004, 02:35:01
Job time : 153.916 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 402.845 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGYHGLYDPLTKLVHMGR.....TDIISVANEDGRRVAAILNH 201

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app_query_fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029020_@cgn_1_1_2156_@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

1	1077	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
2	1077	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
3	1063	98.7	8355	13	US-10-383-201-55	Sequence 55, Appl
4	1063	98.7	8438	13	US-10-042-865-1	Sequence 1, Appli
5	630.5	58.5	12880	16	US-10-295-027-927	Sequence 927, App
6	566	52.6	3614	13	US-10-342-887-1743	Sequence 1743, Ap
7	566	52.6	3614	13	US-10-172-118-1743	Sequence 1743, Ap
8	566	52.6	8473	17	US-10-038-854-39	Sequence 39, Appl
9	566	52.6	8487	17	US-10-038-854-41	Sequence 41, Appl
10	566	52.6	8645	17	US-10-038-854-37	Sequence 37, Appl
11	566	52.6	8675	17	US-10-038-854-35	Sequence 35, Appl
12	556	51.6	791	13	US-09-823-245A-85	Sequence 85, Appl
13	536.5	49.8	8797	9	US-09-808-602-74	Sequence 74, Appl
14	536.5	49.8	8797	9	US-09-808-602-77	Sequence 77, Appl
15	536.5	49.8	8797	10	US-09-800-198-62	Sequence 62, Appl
16	536.5	49.8	8797	10	US-09-800-198-65	Sequence 65, Appl
17	531.5	49.4	8689	9	US-09-808-602-78	Sequence 78, Appl
18	531.5	49.4	8689	10	US-09-800-198-66	Sequence 66, Appl
19	529.5	49.2	3217	15	US-10-198-846-13976	Sequence 13976, A
20	529.5	49.2	6560	9	US-09-808-602-76	Sequence 76, Appl
21	529.5	49.2	6560	10	US-09-800-198-64	Sequence 64, Appl
22	529.5	49.2	9058	16	US-10-144-194A-79	Sequence 79, Appl
23	529.5	49.2	9695	16	US-10-144-194A-81	Sequence 81, Appl
24	529.5	49.2	9729	9	US-09-808-602-12	Sequence 12, Appl
25	529.5	49.2	9729	10	US-09-800-198-12	Sequence 12, Appl
26	529.5	49.2	9826	9	US-09-808-602-7	Sequence 7, Appli
27	529.5	49.2	9826	10	US-09-800-198-7	Sequence 7, Appli
28	525.5	48.8	8409	9	US-09-808-602-79	Sequence 79, Appl
29	525.5	48.8	8409	10	US-09-800-198-67	Sequence 67, Appl
30	520.5	48.3	8575	13	US-10-072-012-143	Sequence 143, App
31	194	18.0	431	9	US-09-954-456-1034	Sequence 1034, Ap
32	188.5	17.5	1973	9	US-09-864-761-4526	Sequence 4526, Ap
33	131.5	12.2	399	9	US-09-861-893-29	Sequence 29, Appl
34	131.5	12.2	399	16	US-10-308-862-6	Sequence 6, Appli
35	128	11.9	464	10	US-09-918-995-27921	Sequence 27921, A
36	93.5	8.7	1860	16	US-10-369-493-41803	Sequence 41803, A
37	88	8.2	1248	16	US-10-260-238-4369	Sequence 4369, Ap
38	87.5	8.1	2836	16	US-10-080-334-59	Sequence 59, Appl
39	86.5	8.0	1157	13	US-10-424-599-96351	Sequence 96351, A
40	86.5	8.0	7158	15	US-10-156-761-201	Sequence 201, App
41	86.5	8.0	9025608	15	US-10-156-761-1	Sequence 1, Appli
42	84.5	7.8	20986	13	US-10-158-844-54	Sequence 54, Appl
43	84	7.8	2419	9	US-09-893-238-8	Sequence 8, Appli
44	84	7.8	2836	16	US-10-320-797-1301	Sequence 1301, Ap
45	84	7.8	4836	16	US-10-320-797-301	Sequence 301, App

Mon Aug 16 09:01:13 2004

us-10-029-020-14 copy 2400 2600.p2n.rnpb

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)

```

```
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 7652 GACCAGCTCTATGGCTCCACAATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAGTTT 7711
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 7712 GCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCTTGAAGGATGCCGAGTG 7771
QY 181 ThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
Db 7772 ACCACAGACATCATCAGTGTGGCCAATGAGGATGGGCGAAGGTTGCTGCCATCTTGAAC 7831
QY 201 His 201
Db 7832 CAT 7834
RESULT 3
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55
Alignment Scores:
Pred. No.: 3.5e-141 Length: 8355
Score: 1063.00 Matches: 201
Percent Similarity: 96.17% Conservative: 0
Best Local Similarity: 96.17% Mismatches: 8
Query Match: 98.70% Indels: 1
DB: 13 Gaps: 1
US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-383-201-55 (1-8355)
QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7192 ATAGGCTACCATGGTGGCTCTATGATCCACTCACAAGCTTGTCACATGGCGCGCGA 7251
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
```

```
Db 7252 GATTATGATGTGCTGGCCGACGCTGGACTAGCCACAGACCACGAGCTGTGGAAGCACCTT 7311
QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7312 AGTAGCAGCAACGTCATGCCTTTTAATCTCTATATGTTCAAAAACAACACCCCATCAGC 7371
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7372 AACTCCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGA 7431
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7432 TTCCAGCTACACAACGTGATCCCTGGTTATCCCAAAACCAGACATGGATGCCATGGAACCC 7491
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys----- 118
Db 7492 TCCTACGAGCTCATCCACACACAGATGAAAACGCGAGGTGGGACACAGCAAGGTAATT 7551
QY 119 -----SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLys 132
Db 7552 CCTGCACAAGGCTGCCAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAG 7611
QY 133 AlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGln 152
Db 7612 GCCTTTGTACCTTAGAACGGTTTGACCAGCTCTATGGCTCCACAATCACCAGCTGCCAG 7671
QY 153 GlnAlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLys 172
Db 7672 CAGGCTCCAAAAGACCAAGAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAG 7731
QY 173 PheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGly 192
Db 7732 TTTGCCTTGAAGGATGGCCGAGTGACCACACAGACATCATCAGTGTGGCCAATGAGGATGG 7791
QY 193 ArgArgValAlaAlaIleLeuAsnHis 201
Db 7792 CGAAGGGTTGCTGCCATCTTGAACCAT 7818
```

```
RESULT 4
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
```

[illegible]

QY 59 IleSerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThr 78
Db 7304 GTTGGCAAAATTCAGATGTTGCAAAGTATACCCACAGACATCAGAAAGTTGGTGAGCTA 7363

QY 79 PheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet 98
Db 7364 TTTGGTTTCCAATTACACAATGTACTACCTGGATTTCCTCCAAACCTGAATTAGAAAAATTTA 7423

QY 99 GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys 118
Db 7424 GAATTAACCTACGAGCTTCTACGGCTTCAGACAAAAAACTCAAGAGTGGGATCCTGGAAAG 7483

QY 119 SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu 138
Db 7484 ACTATCCTGGGATTCAGTGTGAACCTCCAGAAACAGCTCAGGAATTTCAATTCCTTGGAC 7543

QY 139 ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys 158
Db 7544 CAACTACCTATGACTCCCGATACAATGATGGACGGTGCCTTGAAGGAGGGAAGCAACCA 7603

QY 159 LysPheAlaSerSerGlySerValPheGlyLysGlyVallysPheAlaLeuLysAspGly 178
Db 7604 AGGTTTGCTGCTGCCCTTCTGTGTTTGGAAAGGTATAAAATTTGCCATCAAGGATGGC 7663

QY 179 ArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIle 198
Db 7664 ATAGTAACAGCTGATATTATAGGAGTAGCCCAATGAAGATAGCAGCGGCTTGCTGCCATT 7723

QY 199 LeuAsnHis 201
Db 7724 CTCATAAT 7732

RESULT 6
US-10-342-887-1743
; Sequence 1743, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1743
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1743

Alignment Scores:
Pred. No.: 5.07e-70 Length: 3614
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 13 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-342-887-1743 (1-3614)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 2192 ATTGGATTTTCATGGTGGCCTGTATGACCCACTACCAAATTAATCCACTTTGGAGAAAGA 2251

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 2252 GATTATGACATTTTGGCAGGACGGTGGACACACCTGACATAGAAATCTGGAAAAAGAAATT 2311

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 2312 ---GGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCTGCAAGC 2368

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 2369 AAAATCCATGACGCGTGAAGATTACATCACAGATGTTAACAGCTGGTGGTGACATTTGGT 2428

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 2429 TTCCATCTGCACAATGCTATTCTTGGATTCCCTGTTCCCAAATTTGATTTAACAGAACCT 2488

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 2489 TCTTACGAACCTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 2536

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 2537 TTCGGAGTCCAGCAGCAAGTGGCGGCGCAGGCCCAAGGCCTTCCTGCTCGCTGGGGAAGATG 2596

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 2597 GCCGAGGTGCAG-----GTGAGCCGGCGCGCGCGCGCGCGCGAGTCCTGG 2644

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyVallysPheAlaLeuLysAsp 177
Db 2645 CTGTGGTTCGCCACGGTCAAGTCGCTGATCGGCAAGGCGCTCATGTGCGCGTCAGCCAG 2704

QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAla 197
Db 2705 GGCCGGGTGCAGACCACCGTCTCAACATCGCCAACGAGGACTGCATCAAGGTGGCGGCC 2764

QY 198 IleLeuAsnHis 201
Db 2765 GTGCTCAACAC 2776

RESULT 7
US-10-172-118-1743
; Sequence 1743, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1743
; LENGTH: 3614
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_018104
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1743

```
Alignment Scores:
Pred. No.: 5.07e-70 Length: 3614
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 13 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-172-118-1743 (1-3614)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 2192 ATTGGATTTTCATGGTGGCCCTGTATGACCCACTCACCAAAATTAATCCACTTTGGAGAAAGA 2251

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 2252 GATTATGACATTTTGGCAGGACGGTGGACAACACCTGCATAGAAATCTGGAAAGAAT 2311

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 2312 ---GGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCCCTGCAAGC 2368

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 2369 AAAATCCATGACGTGAAGATTACATCACAGATGTTAAACAGCTGGCTGGTGACATTTGGT 2428

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 2429 TTCCATCTGCACAATGCTATTCTCTGGATTCCTGTTCCCAAAATTGATTAAACAGAACCT 2488

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 2489 TCTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 2536

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 2537 TTCGGAGTCCAGCAGCAAGTGGCGGAGGCCAAGGCCTTCCTGTCGCTGGGAAGATG 2596

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 2597 GCCAGGTGCAG-----GTGAGCCGGCGCGCGCGCGCGAGTCTCTGG 2644

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 2645 CTGTGTTCCGCACGGTCAAGTCGTCATCGCAAGGGCGTCTGCTGCCGTCAGCCAG 2704

QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
Db 2705 GGCGCGGTGCAGACCAACGTCGTCAACATCGCCAACGAGGAGTGCATCAAGTGGCGGCC 2764

QY 198 IleLeuAsnHis 201
Db 2765 GTGCTCAACAAC 2776

RESULT 8
US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
```

```
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-038-854-39

Alignment Scores:
Pred. No.: 2.03e-69 Length: 8473
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 17 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-038-854-39 (1-8473)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7050 ATTGGATTTTCATGGTGGCCCTGTATGACCCACTCACCAAAATTAATCCACTTTGGAGAAAGA 7109

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 7110 GATTATGACATTTTGGCAGGACGGTGGACAACACCTGCATAGAAATCTGGAAAGAAT 7169

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 7170 ---GGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCCCTGCAAGC 7226

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7227 AAAATCCATGACGTGAAGATTACATCACAGATGTTAAACAGCTGGCTGGTGACATTTGGT 7286

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
```

Db 7287 TTCCATCTGCACAATGCTATTCTCGGATTCCTGTCCCAAATTTGATTAAACAGAACCT 7346
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7347 TCTTACGAACCTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7394
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7395 TTCGGAGTCCAGCAGCAAGTGGCGGCAGGCCAAGGCCTTCTGTGCTGGGAAGATG 7454
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7455 GCCGAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCGCAGTCCTGG 7502
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7503 CTGTGGTTCCGCACGGTCAAGTCGTGATCGGCAAGGGCGTCATGCTGCGCGTCAGCCAG 7562
QY 178 GlyArgValThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgValAlaAla 197
Db 7563 GGCGCGCTGCAGACCAACGTGCTCAACATCGCCAAACGAGGACTGTCATCAAGGTGGCGGCC 7622
QY 198 IleLeuAsnHis 201
Db 7623 GTGCTCAACAAC 7634

RESULT 9

US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernnet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 2,04e-69 Length: 8487
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 17 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-038-854-41 (1-8487)
QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7046 ATTGGATTTTCATGGTGGCCTGTATGACCCACTCACCAATAATTAATCCACTTTGGAGAAAGA 7105
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7106 GATTATGACATTTTGGCAGGACGGTGGACAACACCTGACATAGAAATCTGGAAAAGAAAT 7165
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnAsnProIleSer 60
Db 7166 ---GGGAAGGACCCAGCTCTTTTAACTTGACATGTTTAGGAATAACAACCCCTGCAAGC 7222
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7223 AAAATCCATGACGTGAAGATTACATCACAGATGTTAACAGCTGGTGGTGGTGGTGGT 7282
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7283 TTCCATCTGCACAATGCTATTCTTGGATTCTTCTGTTCCAAATTTGATTAAACAGAACCT 7342
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7343 TCTTACGAACCTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7390
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7391 TTCGGAGTCCAGCAGCAAGTGGCGGCAGGCCCAAGGCCTTCTGTGCTGGGGAAGATG 7450
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7451 GCCGAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCGCAGTCCTGG 7498
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7499 CTGTGGTTCCGCACGGTCAAGTCGTGATCGGCAAGGGCGTCATGCTGCGCGCTCAGCCAG 7558
QY 178 GlyArgValThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgValAlaAla 197
Db 7559 GGCGCGCTGCAGACCAACGTGCTCAACATCGCCAAACGAGGACTGTCATCAAGGTGGCGGCC 7618
QY 198 IleLeuAsnHis 201
Db 7619 GTGCTCAACAAC 7630

; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:
Pred. No.: 2.11e-69 Length: 8675
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 17 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-10-038-854-35 (1-8675)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7234 ATTGGATTTCATGGTGGCCTGTATGACCCACTCACCAATTAATCCACTTTGGAGAAAGA 7293

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7294 GATTATGACATTTTGGCAGGACGGTGGACACACCTGACATAGAAATCTGGAAAAGAAATT 7353

QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7354 ---GGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCCCTGCAAGC 7410

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7411 AAAATCCATGACGTGAAAGATTATACATCACAGATGTTAAACAGCTGGCTGTGACATTTGGT 7470

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7471 TTCCATCTGCACAATGCTATTTCCTGGATTCCCTGTTCCCAAAATTGAATTAACAGAACCT 7530

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7531 TCTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7578

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140

Db 7579 TTCGGAGTCCAGCAGCAAGTGGCGGGCAGGCCAAGGCCTTCCTGTCTGGGGAAGATG 7638
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7639 GCCGAGGTGCAG-----GTGAGCGCGCGCGCGCGCGCGCGCGCGCAGTCCTGG 7686
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7687 CTGTGTTTCCGCACCGTCAAGTCGCTGATCGGCAAGGCGCTCATGCTGCGCGTCAGCCAG 7746
QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAla 197
Db 7747 GGCCGCGTGCAGACCAACGTTGCTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGGCC 7806
QY 198 IleLeuAsnHis 201
Db 7807 GTGCTCAACAAC 7818

RESULT 12

US-09-823-245A-85
; Sequence 85, Application US/09823245A
; Publication No. US20020039760A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6401
; CURRENT APPLICATION NUMBER: US/09/823,245A
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/194,941
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 85
; LENGTH: 791
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 566,604,734,779
; OTHER INFORMATION: n = a,t,g, or c
US-09-823-245A-85

Alignment Scores:
Pred. No.: 1.18e-69 Length: 791
Score: 556.00 Matches: 106
Percent Similarity: 71.14% Conservative: 37
Best Local Similarity: 52.74% Mismatches: 52
Query Match: 51.62% Indels: 7
DB: 13 Gaps: 3

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-823-245A-85 (1-791)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 102 ATTGGATTTCATGGTGGCCTGTATGACCCACTCACCAATTAATCCACTTTGGAGAAAGA 161

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 162 GATTATGACATTTTGGCAGGACGGTGGACACACCTGACATAGAAATCTGGAAAAGAAATT 221

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 222 ---GGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCCCTGCAAGC 278

QY	61	AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly	80
Db	279	AAAATCCATGACGTGAAGATTACATCAGATGTTAACAGCTGGCTGGTGACATTGGT	338
QY	81	PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro	100
Db	339	TTCCATCTGCACAATGCTATTCTTGATTCTCTGCTTCCAAATTTGATTAAACAGAACCT	398
QY	101	SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle	120
Db	399	TCTTACGAACCTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGTCCATC	446
QY	121	LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe	140
Db	447	TTCGGAGTCCAGCAGCAAGTGGCGCGCAGGCCAAGGCTCTCTGCTGCGTGGGGAAGATG	506
QY	141	AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe	160
Db	507	GCCGAGGTGACGGTGAGCCGGCGCGCGCGCGCAGTCCTGGCTGTGG---TTC	562
QY	161	AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal	180
Db	563	GCCNCGGTCAAGTCGTGATCGGCAAGGCGTCATGTGTCGTCAGCCAGGGCGCGGTG	622
QY	181	ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn	200
Db	623	CAGACCAACGTGCTCAACATCGCCAAACGAGGACTGCATCAAGGTGGCGCGCTCAAC	682
QY	201	His 201	
Db	683	AAC 685	
RESULT 13			
US-09-808-602-74			
; Sequence 74, Application US/09808602			
; Patent No. US20020155115A1			
; GENERAL INFORMATION:			
; APPLICANT: Vernet, Corine A			
; APPLICANT: Fernandes, Elma			
; APPLICANT: Shimkets, Richard A			
; APPLICANT: Herrman, John L			
; APPLICANT: Majumder, Kumud			
; APPLICANT: Mishra, Vishnu			
; APPLICANT: Mezes, Peter S			
; APPLICANT: MacDougall, John			
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same			
; FILE REFERENCE: 15966-697 CIP			
; CURRENT APPLICATION NUMBER: US/09/808,602			
; CURRENT FILING DATE: 2001-03-14			
; PRIOR APPLICATION NUMBER: 09/800,198			
; PRIOR FILING DATE: 2001-03-05			
; PRIOR APPLICATION NUMBER: 60/186,596			
; PRIOR FILING DATE: 2000-03-03			
; NUMBER OF SEQ ID NOS: 114			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 74			
; LENGTH: 8797			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
US-09-808-602-74			
Alignment Scores:			
Pred. No.:	3.93e-65	Length:	8797
Score:	536.50	Matches:	104
Percent Similarity:	68.78%	Conservative:	37
Best Local Similarity:	50.73%	Mismatches:	49
Query Match:	49.81%	Indels:	15
DB:	9	Gaps:	4
US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-808-602-74 (1-8797)			
QY	1	IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg	20
		:::	

Db	7307	ATTGGCTTCCAGGAGGCCTCTATGACCCCTCACCAAGCTCGTCCACTTTACTCAACGT	7366
QY	21	AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu	40
Db	7367	GATTATGACGTGCTGGCAGGACGGTGGAGCTCCCCGACTACACCATGTGGAGAACGTG	7426
QY	41	SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer	60
Db	7427	GGCAAGGAGCCAGCC---CCCTTCAACCTGTACATGTTCAAGAACAAACATCTCTGAGC	7483
QY	61	AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly	80
Db	7484	AATGAGCTGGACTTAAAGAACTACGTGACAGACGTGAAGAGCTGGCTTGTGATGTTTGA	7543
QY	81	PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro	100
Db	7544	TTTCAGCTCAGCAACATCATCTCTGGATTCCCGAGAGCCAAATGTATTTTGTGCTTCCC	7603
QY	101	SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle	120
Db	7604	CCCTATGAACCTGTACAGAGTCAAGCAAGC-----GAGAACGGACAGCTCATT	7651
QY	121	LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe	140
Db	7652	ACAGGTGTCCAGCAGCAACTGAGAGGCATAACCCAGGCTTCTTGGCTCTGGAA-----	7705
QY	141	AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys---	159
Db	7706	-----GGACAGGTGCATCACTATAAAGCTCCATGCCAGCATCCGAGAGAAAGCA	7753
QY	160	-----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys	176
Db	7754	GGCCACTGGTTTGCTACCAACACACCCCATCATCGGCAAGGCATCATGTTTGGCATCAA	7813
QY	177	AspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAla	196
Db	7814	GAAGGCGGGTGACCCACAGGACTGTCTAGCATCGCCAGTGAGGACAGCCGCAAGGTAGCA	7873
QY	197	AlaIleLeuAsnHis 201	
Db	7874	TCCGTGTGAACAAT 7888	
RESULT 14			
US-09-808-602-77			
; Sequence 77, Application US/09808602			
; Patent No. US20020155115A1			
; GENERAL INFORMATION:			
; APPLICANT: Vernet, Corine A			
; APPLICANT: Fernandes, Elma			
; APPLICANT: Shimkets, Richard A			
; APPLICANT: Herrman, John L			
; APPLICANT: Majumder, Kumud			
; APPLICANT: Mishra, Vishnu			
; APPLICANT: Mezes, Peter S			
; APPLICANT: MacDougall, John			
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same			
; FILE REFERENCE: 15966-697 CIP			
; CURRENT APPLICATION NUMBER: US/09/808,602			
; CURRENT FILING DATE: 2001-03-14			
; PRIOR APPLICATION NUMBER: 09/800,198			
; PRIOR FILING DATE: 2001-03-05			
; PRIOR APPLICATION NUMBER: 60/186,596			
; PRIOR FILING DATE: 2000-03-03			
; NUMBER OF SEQ ID NOS: 114			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 77			
; LENGTH: 8797			
; TYPE: DNA			
; ORGANISM: Mus musculus			
US-09-808-602-77			
Alignment Scores:			
Pred. No.:	3.93e-65	Length:	8797


```
Score: 536.50 Matches: 104
Percent Similarity: 68.78% Conservative: 37
Best Local Similarity: 50.73% Mismatches: 49
Query Match: 49.81% Indels: 15
DB: 9 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-808-602-77 (1-8797)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7307 ATTGGCTTCCACGAGGCGCTCTATGACCCCTCACCAGCTCGTCCACTTTACTCAACGT 7366

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7367 GATTATGACGTGCTGGCAGGACGGTGGACGTCCCCGACTACACCATGTGGAGGAACGTG 7426

QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7427 GGCAAGGAGGCCAGCC---CCCTTCAACCTGTACATGTTCAAGAACAAATCCTCTGAGC 7483

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThrPheGly 80
Db 7484 AATGAGCTGGACTTAAAGAACTACGTGACAGACGTTGAAGAGCTGGCTTGTGATTTTGA 7543

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7544 TTTCAGCTCAGCAACATCATTCCTGGATTCCCGAGAGCCAAATGATTTTGTGCCTCCC 7603

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 7604 CCTATGAACCTGCAGAGAGTCAAGCAAGC-----GAGAACGGACAGCTCATT 7651

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7652 ACAGGTGTCAGCAGACAACTGAGAGGCATAACCAAGGCCTTCTGGCTCTGGAA----- 7705

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7706 -----GGACAGTGCATCACTAAAAAGCTCCATGCCAGCATCCGAGAGAAAGCA 7753

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys 176
Db 7754 GGCCACTGGTTTGTACCAACACACCCCATCATCGGCAAAAGSCATCATGTTTGCCATCAA 7813

QY 177 AspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAla 196
Db 7814 GAAGGGCGGGTGACCACAGGAGTGTCTAGCATCGCCAGTGAGGACAGCCGCAAGGTAGCA 7873

QY 197 AlaIleLeuAsnHis 201
Db 7874 TCCGTGTTGAACAAT 7888
```

RESULT 15

```
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62
```

Alignment Scores:

```
Pred. No.: 3.93e-65 Length: 8797
Score: 536.50 Matches: 104
Percent Similarity: 68.78% Conservative: 37
Best Local Similarity: 50.73% Mismatches: 49
Query Match: 49.81% Indels: 15
DB: 10 Gaps: 4
```

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-800-198-62 (1-8797)

```
QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7307 ATTGGCTTCCACGAGGCGCTCTATGACCCCTCACCAGCTCGTCCACTTTACTCAACGT 7366

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7367 GATTATGACGTGCTGGCAGGACGGTGGACGTCCCCGACTACACCATGTGGAGGAACGTG 7426

QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7427 GGCAAGGAGGCCAGCC---CCCTTCAACCTGTACATGTTCAAGAACAAATCCTCTGAGC 7483

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThrPheGly 80
Db 7484 AATGAGCTGGACTTAAAGAACTACGTGACAGACGTTGAAGAGCTGGCTTGTGATTTTGA 7543

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7544 TTTTCAGCTCAGCAACATCATTCCTGGATTCCCGAGAGCCAAATGATTTTGTGCCTCCC 7603

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 7604 CCTATGAACCTGCAGAGAGTCAAGCAAGC-----GAGAACGGACAGCTCATT 7651

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7652 ACAGGTGTCAGCAGACAACTGAGAGGCATAACCAAGGCCTTCTGGCTCTGGAA----- 7705

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7706 -----GGACAGTGCATCACTAAAAAGCTCCATGCCAGCATCCGAGAGAAAGCA 7753

QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys 176
Db 7754 GGCCACTGGTTTGTACCAACACACCCCATCATCGGCAAAAGGCATCATGTTTGCCATCAA 7813

QY 177 AspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAla 196
Db 7814 GAAGGGCGGGTGACCACAGGAGTGTCTAGCATCGCCAGTGAGGACAGCCGCAAGGTAGCA 7873

QY 197 AlaIleLeuAsnHis 201
Db 7874 TCCGTGTTGAACAAT 7888
```

Search completed: August 14, 2004, 19:22:42
Job time : 457.845 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 59.4656 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGYHGLYDPLTKLVHMGR.....TDIISVANEDGRRVAAILNH 201

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues 1365418
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1 1 258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/pctus_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	84.5	7.8	20986	4	US-08-961-527-54
2	84	7.8	2419	3	US-09-245-041-8
3	84	7.8	8827	3	US-09-245-041-1
4	82.5	7.7	1650	4	US-09-543-681A-548
5	81	7.5	45175	4	US-09-453-702B-116
C 6	81	7.5	49795	4	US-09-453-702B-60
7	80	7.4	2280	4	US-09-328-352-115
C 8	80	7.4	9909	4	US-08-961-527-12
9	78.5	7.3	3215	4	US-09-620-312D-86
10	78	7.2	3254	1	US-08-162-809-15
11	77.5	7.2	5706	4	US-09-738-946-11
12	76.5	7.1	1119	4	US-09-252-991A-2392

13	76.5	7.1	2772	4	US-09-717-926-3	Sequence 3, Appli
14	76.5	7.1	7220	4	US-09-717-926-1	Sequence 1, Appli
15	76	7.1	2745	4	US-09-817-514A-5	Sequence 5, Appli
16	75.5	7.0	3132	4	US-08-851-567B-60	Sequence 60, Appli
17	75	7.0	41708	4	US-09-470-512A-3	Sequence 3, Appli
C 18	74.5	6.9	1044	3	US-09-202-832-2	Sequence 2, Appli
C 19	74.5	6.9	1044	3	US-09-202-832-11	Sequence 11, Appli
C 20	74.5	6.9	1247	3	US-09-202-832-3	Sequence 3, Appli
21	74.5	6.9	1533	4	US-09-134-000C-3230	Sequence 3230, Ap
C 22	74.5	6.9	1873	4	US-09-016-434-1437	Sequence 1437, Ap
23	74.5	6.9	2397	4	US-09-134-001C-2080	Sequence 2080, Ap
24	74.5	6.9	3852	4	US-09-245-248B-29	Sequence 29, Appli
C 25	74	6.9	11823	4	US-08-956-171E-136	Sequence 136, App
C 26	73.5	6.8	1309	1	US-07-926-788A-1	Sequence 1, Appli
27	73.5	6.8	2625	3	US-09-245-041-18	Sequence 18, Appli
28	73.5	6.8	2895	4	US-09-543-681A-863	Sequence 863, App
29	73.5	6.8	4072	3	US-09-245-041-16	Sequence 16, Appli
30	73.5	6.8	8589	3	US-09-245-041-14	Sequence 14, Appli
31	73	6.8	720	4	US-09-252-991A-13392	Sequence 13392, A
32	73	6.8	1296	4	US-09-527-058-5	Sequence 5, Appli
33	73	6.8	3540	4	US-09-107-532A-2472	Sequence 2472, Ap
34	73	6.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
35	72.5	6.7	849	4	US-09-134-000C-570	Sequence 570, App
36	72.5	6.7	4849	2	US-08-540-804-13	Sequence 13, Appli
37	72.5	6.7	4849	2	US-08-218-265-13	Sequence 13, Appli
38	72.5	6.7	4849	3	US-08-521-872-13	Sequence 13, Appli
39	72.5	6.7	4849	3	US-08-590-399-13	Sequence 13, Appli
40	72.5	6.7	37948	3	US-09-251-645-11	Sequence 11, Appli
C 41	72.5	6.7	118067	4	US-09-497-855A-32	Sequence 32, Appli
42	72	6.7	1347	4	US-09-134-001C-1309	Sequence 1309, Ap
43	72	6.7	2202	4	US-08-956-171E-284	Sequence 284, App
C 44	72	6.7	2329	1	US-08-253-785-2	Sequence 2, Appli
45	72	6.7	2920	4	US-08-976-259-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-08-961-527-54/c
; Sequence 54, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:


```

; LENGTH: 20986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-54

Alignment Scores:
Pred. No.: 24.1 Length: 20986
Score: 84.50 Matches: 31
Percent Similarity: 43.01% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 38
Query Match: 7.85% Indels: 15
DB: 4 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x US-08-961-527-54 (1-20986)

QY 34 HisGluLeuTrpLysHisLeuSerSerAsnValMetPropPheAsnLeuTyrMetPhe 53
Db 13300 CATACAAGAGGAAGCATCTACTATCCAAG-----TTC 13268

QY 54 LysAsn-----AsnAsnProIleSerAsnSerGlnAspIleLysCysPheMetThrAsp 71
Db 13267 AAGAAATATCTTAGGAATGACATTTCGAATTCTCTTCT-----TTTTTAACAGCA 13217

QY 72 ValAsnSerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrPro 91
Db 13216 CTGACTTTCTTAAACAAGACGTTTGTGTTTGGTTTACAAAATCTAGTTCCCTATTATAAT 13157

QY 92 LysProAspMetAspAlaMetGluProSerTyrGluLeuIleHis---ThrGlnMetLys 110
Db 13156 TATTACGAACATTATGGGATTGAAAGCGACAGCCACTGATCATATTTTCGAAAGCAATT 13097

QY 111 ThrGlnGluTrpAspAsnSerLysSerIleLeuGlyVal 123
Db 13096 GTTCAAGAGTGGATGACTGAGCAGAAGATTGAAGGAGTT 13058

RESULT 2
US-09-245-041-8
; Sequence 8, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-8

Alignment Scores:
Pred. No.: 0.956 Length: 2419
Score: 84.00 Matches: 54
Percent Similarity: 37.67% Conservative: 30
Best Local Similarity: 24.22% Mismatches: 86
Query Match: 7.80% Indels: 53
DB: 3 Gaps: 10

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-245-041-8 (1-2419)

QY 2 GlyTyr---HisGlyLeuTyrAspProLeuThrLys-----LeuValHisMetGly 18
Db 1764 GGTATGGCCACAGTAGTGTATTATGATGACAGGACCAAGGCTCTGTACGTTTCATGGTGGC 1823
```

```

QY 19 ArgArg-----AspTyrAspVal 24
Db 1824 TACAAGGCTTTCAGCGCCAAATAACCGGCTTGCAGATGACCTCTACAGATACGATGTG 1883

QY 25 LeuAlaGlyArgTrpThr---SerProAspHisGluLeuTrpLysHisLeuSerSer 43
Db 1884 GATACTCAGATGTGGACCATTTCTTAAGACAGCCGATTTTTCGGTTACTTGCATACAGCT 1943

QY 44 AsnValMetProPheAsnLeuTyrMetPheLysAsnAsn-----AsnProIleSerAsn 61
Db 1944 GTGATAGTAGTGGAACCATGCTGTTGGAGGGAACACACACAATGACACTTCCATG 2003

QY 62 SerGlnAspIleLysCysPheMetThrAsp-----ValAsnSer 74
Db 2004 AGCCACGGTGCCAAATGCTTCTCTCGGACTTCATGGCTTATGACATTGCTTGTGACCGA 2063

QY 75 TrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAsp 94
Db 2064 TGGTCAGTGTCTCCAGACCTGAGCTCCATCAT-----GAT 2099

QY 95 MetAspAlaMetGluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrp 114
Db 2100 GTCAACAGATTTGGCCATTTCAGCAGTCTTGTACAAACAGCACCATTGATGTGTCGGCGGC 2159

QY 115 AspAsnSerLysSerIleLeuGlyVal-----GlnCysGluValGln 128
Db 2160 TTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGGGATGCACAC 2219

QY 129 LysGlnLeuLysAlaPheValThr-----LeuGluArgPheAspGlnLeuTyr 144
Db 2220 CGCAGTGAAGCTGCTTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAG 2279

QY 145 GlySerThrIleThrSerCysGlnGlnAlaPro-LysThrLysLysPheAlaSerSerG1 164
Db 2280 TCGTCTCGATGTACCTCCTCGGAGTTGGCAACTGAAGAACAGCAGAAAGTTAAATCA 2339

QY 164 ySerValPheGlyLysGlyVallyLysPheAlaLeuLysAspGlyArgValThrThrAsp11 184
Db 2340 GAGTGTTTT-----CTAAAGAAACCCCTTGACCATGACAGATGT 2378

QY 184 elleSer 186
Db 2379 GACCAGC 2385

RESULT 3
US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-1

Alignment Scores:
Pred. No.: 7.24 Length: 8827
Score: 84.00 Matches: 54
Percent Similarity: 37.67% Conservative: 30
```

Best Local Similarity: 24.22% Mismatches: 86
Query Match: 7.80% Indels: 53
DB: 3 Gaps: 10

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-245-041-1 (1-8827)

```
QY      2  GlyTyr---HisGlyGlyLeuTyrAspProLeuThrLys-----LeuValHisMetGly 18
      ||||| ||| :::||||| ||||| ||||| ||||| |||||
Db     1764 GGTATGGCCACAGTAGTGTATTATGATGACAGGACCAAGGCTCTGTACGTTTCATGGTGGC 1823

QY     19  ArgArg-----AspTyrAspVal 24
      ::|
Db    1824 TACAAGGCTTTCAGCGCCCAACAATACCGGCTTGCAGATGACCTCTACAGATACGATGTG 1883

QY     25  LeuAlaGlyArgTrpThr---SerProAspHisGluLeuTyrLysHisLeuSerSerSer 43
      ||||| ||| :::||||| ||||| ||||| ||||| |||||
Db    1884 GATACTCAGATGTGGACCAATTCTTAAGACAGCCGATTTTCCGTACTTGCATACAGCT 1943

QY     44  AsnValMetProPheAsnLeuTyrMetPheLysAsnAsn-----AsnProIleSerAsn 61
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    1944 GTGATAGTGAGTGAACCATGCTGCTCTCCGACTTCATGGCTTATGACATGCTTGTGACCGA 2003

QY     62  SerGlnAspIleLysCysPheMetThrAsp-----ValAsnSer 74
      ||||| ||||| ||||| ||||| ::|
Db    2004 AGCCACGGTGCCAAATGCTTCTCTCCGACTTCATGGCTTATGACATGCTTGTGACCGA 2063

QY     75  TrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAsp 94
      ||||| ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    2064 TGGTCAGTGCTTCCCAGACCTGAGCTCCATCAT-----GAT 2099

QY     95  MetAspAlaMetGluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrp 114
      ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    2100 GTCAACAGATTGGGCCATTGAGCAGTCTTGTACAACAGCACCATGATGTTGCGCGGC 2159

QY    115  AspAsnSerLysSerIleLeuGlyVal-----GlnCysGluValGln 128
      ||||| ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    2160 TTCAACAGCCTCCTCCTCAGTGACGTCTTGGTCTTTACCTCGGAGCAGTGCATGCACAC 2219

QY    129  LysGlnLeuLysAlaPheValThr-----LeuGluArgPheAspGlnLeuTyr 144
      ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    2220 CGCAGTGAAGTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAG 2279

QY    145  GlySerThrIleThrSerCysGlnGlnAlaPro-LysThrLysLysPheAlaSerSerG1 164
      ||||| ::|::|::|::|::|::|::|::|::|::|::|::|::|
Db    2280 TCGTCTCGATGTACCTCCTGGGAGTTGGCAACTGAGAACAAAGCAGAAAAGTTAAATCA 2339

QY    164  ySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgValThrThrAspI1 184
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    2340 GAGTGTGTTTT-----CTAAAGAACCCCTTGACCATGACAGATGT 2378

QY    184  eileSer 186
      : |||
Db    2379 GACCAGC 2385
```

RESULT 4
US-09-543-681A-548
; Sequence 548, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 548
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-548

Mon Aug 16 09:01:12 2004

us-10-029-020-14_copy_2400_2600.p2n.rn1

FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 116:
SEQUENCE CHARACTERISTICS:
LENGTH: 45175
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 116:
US-09-453-702B-116

Alignment Scores:
Pred. No.: 233 Length: 45175
Score: 81.00 Matches: 44
Percent Similarity: 39.25% Conservative: 29
Best Local Similarity: 23.66% Mismatches: 75
Query Match: 7.52% Indels: 38
DB: 4 Gaps: 9

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-453-702B-116 (1-45175)

QY 23 AspValLeuAlaGly-----ArgTrpThrSer----- 31
Db 42089 GATGTGATGCAGGAACAATTGACGGTACCATCCATCGCTGGCGCAAAATGAGGTTATC 42148
QY 32 ---ProAspHisGluLeuTrpLysHisLeuSerSerSer-----AsnVal 45
Db 42149 GACAAGGATCGCGAAGACTGGAGCGGTATCTCAGCATCAATGCGCAACAGCCCGAAGCA 42208
QY 46 MetProPheAsn-----LeuTyrMetPheLysAsnAsnProIleSerAsnSer 62
Db 42209 CTTGGCTATGACCGTCAGACTATTTTGGCCCTTGTTGCGGACCGTCCGATCGATATTCAC 42268
QY 63 GlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGln 82
Db 42269 AAAGATCCCGTTGCACTGAACAAATATATCAGTGAATACCTGACGACAAAGGGCGTGT 42328
QY 83 LeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluProSerTyr 102
Db 42329 GAACATGAAGAAACA-----GACCAGAGCTCTACTGATGCTCTCCAGCCGTCAGCA 42379
QY 103 GluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuGly 122
Db 42380 GCACAAACT---GCTCCAGTGGAGACGCGCAGAAATCCGATACCTCAAAAAAATGAAATCCTG 42436
QY 123 ValGlnCysGluValGlnLysGlnLeuLysAla-----PheValThrLeuGluArg 139
Db 42437 GTGGAAGCTGAACCATCTGTAGACGCTGAAGGACCATTTTATTTCGCTCTTTACCGATAAG 42496
QY 140 PheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLys----- 156
Db 42497 GGCGGGGAAATAACGGC-----AGGGCAACAACTTTCTGTT 42535
QY 157 ---ThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeu 175
Db 42536 CTGGACAAGCGCTGGCTGCCGGCGGTACCGGAATCTCAAAGAGAATATTTTGCCCGA 42595
QY 176 LysAspGlyArgValThr 181
Db 42596 AAAAATGGCACATACACG 42613

RESULT 6

US-09-453-702B-60/c
Sequence 60, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 49795
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-453-702B-60

Alignment Scores:
Pred. No.: 271 Length: 49795
Score: 81.00 Matches: 44
Percent Similarity: 39.25% Conservative: 29
Best Local Similarity: 23.66% Mismatches: 75
Query Match: 7.52% Indels: 38
DB: 4 Gaps: 9

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-453-702B-60 (1-49795)
QY 23 AspValLeuAlaGly-----ArgTrpThrSer----- 31
Db 3137 GATGTGATGCAGGAACAATTGACGGTACCATCCATCGCTGGCGCAAAATGAGGTTATC 3078
QY 32 ---ProAspHisGluLeuTrpLysHisLeuSerSerSer-----AsnVal 45
Db 3077 GACAAGGATCGCGAAGACTGGAAGCGGTATCTCAGCATCAATGCGCAACAGCCCGAAGCA 3018
QY 46 MetProPheAsn-----LeuTyrMetPheLysAsnAsnProIleSerAsnSer 62
Db 3017 CTTGGCTATGACCGTCAGACTATTTTGGCCCTTGTTGCGAACGTCCTCGATATTCAC 2958
QY 63 GlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGln 82
Db 2957 AAAGATCCCGTTGCACTGAACAAATATATCAGTGAATACCTGACGACAAAGGGCGTGT 2898

QY 83 LeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluProSerTyr 102
Db 2897 GAACATGAAGAAACA-----GACCAGAGCTCTACTGATGCTCTCCAGCCGTCAGCA 2847
QY 103 GluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuGly 122
Db 2846 GCACAAACT---GTCACAGTGGAGACGGCAGAAATCCGATACTCAAAATAATGAAATCCTG 2790
QY 123 ValGlnCysGluValGlnLysGlnLeuLysAla-----PheValThrLeuGluArg 139
Db 2789 GTGGAAGCTGAACCATCTGTAGAGCGTGAAGGACCATTTATTTCGTCTTTACCGATAAG 2730
QY 140 PheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLys----- 156
Db 2729 GCGGGGGAATAATACGGC-----AGGGCAAAACAAACTTTCTTCTGGT 2691
QY 157 ---ThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeu 175
Db 2690 CTGGACAAGCGCTGGCTGCCGGCGGTACCGAAATCTCAAAAGAAGAAATATTTTGCCCGA 2631
QY 176 LysAspGlyArgValThr 181
Db 2630 AAAAATGGCACATACG 2613
RESULT 7
US-09-328-352-115
; Sequence 115, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 115
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (2238)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkno
US-09-328-352-115
Alignment Scores:
Pred. No.: 2.95 Length: 2280
Score: 80.00 Matches: 27
Percent Similarity: 43.33% Conservative: 12
Best Local Similarity: 30.00% Mismatches: 41
Query Match: 7.43% Indels: 10
DB: 4 Gaps: 3
US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-328-352-115 (1-2280)
QY 116 AsnSerLysSerIleLeuGly-----ValGlnCysGluValGlnLysGln 130
Db 1564 AATAGCAAGCACGCAACAGGTAATGCATTTGACTTCACGCTAGATGATCGCAAAAGTCT 1623
QY 131 LeuLysAlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSer 150
Db 1624 GGTGAGGCTGTTACACAACCTTGAGCAAAATGGCTAAAGATACGGCTTTGTGATTCGCGTT 1683
QY 151 CysGlnGlnAlaProLysThrLysLysPheAlaSerSerGly-----SerVal 166
Db 1684 CTTGATGAGTACAGAAGAAGTCTGCTCGTGCACCGGGGGGCATATTTCATGTATCTGTG 1743
QY 167 PheGly---LysGlyValLysPheAlaLeuLysAspGlyArgValThrThrAspIleIle 185
Db 1744 CTTGGCTACAAAGGCACAGCAGATGCATTAAAGATGCAAAATGCAGAGCTGTGATTGTT 1803
QY 186 SerValAlaAsnGluAspGlyArgArgVal 195

Db 1804 CAGAAAGCAAATGATGAAGCTACAAAAATT 1833
RESULT 8
US-08-961-527-12/c
; Sequence 12, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-12
Alignment Scores:
Pred. No.: 29.4 Length: 9909
Score: 80.00 Matches: 35
Percent Similarity: 40.37% Conservative: 30
Best Local Similarity: 21.74% Mismatches: 72
Query Match: 7.43% Indels: 24
DB: 4 Gaps: 7
US-10-029-020-14_COPY_2400_2600 (1-201) x US-08-961-527-12 (1-9909)
QY 59 IleSerAsn---SerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeu--- 76
Db 3313 GTTTCCAATGTTAGCCATGAGTTACGGACTCCTCTGACTAGCGTAAATCCTATCTTGAA 3254
QY 77 ---LeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMet 95
Db 3253 GCCTTGGATGAGGGGGCTTTGTGTGAAACTGTAGCACCAGACTTTATCAAGGTTTCTCTT 3194
QY 96 AspAlaMetGluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAsp 115
Db 3193 GATGAGACCAACCGTATGATGCGCATGGTGACGGATCTCCTCCATCTTTCACGTATTGAT 3134
QY 116 AsnSerLysSerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheVal 135
Db 3133 AATGCTACCACTCACCTAGATGTGGAACGTG-----ATTACTTCACTGCTTTTATT 3083
QY 136 Thr-----LeuGluArgPheAspGlnLeuTyrGlySer----- 146

Db 3082 ACCTTTATCCTCAATCGTTTTCAGCAAGATGAAAGGACAGGAAAGGAGAAAAAATATGAG 3023
QY 147 -----ThrIleThrSerCysGlnGlnAlaProLysThrLysLysPheAla 161
Db 3022 TTGGTGAGAGATTATCCCATCAATTCTATCTGATGGAAATTGATACAGATAAGATGACG 2963
QY 162 Ser---SerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 2962 CAGGTTGTCGACATATTTTAAATAATGCTATTAAGTATTTCGCCAGATGGGGGTAAATC 2903
QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsn 200
Db 2902 ACT-----GTCAGATGAAGACAACTGAAGACCAGATGATTTTATCCATTTCTGAC 2852
QY 201 His 201
Db 2851 CAC 2849

RESULT 9
US-09-620-312D-86
; Sequence 86, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 86
; LENGTH: 3215
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (264)..(3128)
US-09-620-312D-86

Alignment Scores:
Pred. No.: 7.99 Length: 3215
Score: 78.50 Matches: 48
Percent Similarity: 35.53% Conservative: 33
Best Local Similarity: 21.05% Mismatches: 74
Query Match: 7.29% Indels: 73
DB: 4 Gaps: 10

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-620-312D-86 (1-3215)

QY 1 IleGlyTyrHisGlyGlyLeu-----TyrAsp 9
Db 1017 ATGGGTACAATGCGGAACCTCATTCAGTGTTTAAGACCCTGCCAGCAAGATATGAT 1076

QY 10 ProLeuThrLysLeuValHisMetGlyArgAspTyrAspValLeuAlaGlyArgTrp 29
Db 1077 CCTGACACCAAGACGTTGGAACCTTCAGCATGAATGACTATAGTCCCTGATG-----1127
QY 30 ThrSerProAspHisGluLeuTrpLysHisLeuSerSerAsnValMetPropPheAsn 49
Db 1128 -----AAAGCAGCCCAGAGCCTCCCCACGGTCAACCTGCAGCCTCTGGAA 1172
QY 50 LeuTyrMetPheLysAsnAsnProIleSerAsnSerGlnAspIle-----65
Db 1173 TGGGCTATGGCAGCAGGAGTCAACCTCCACCAGCAGTGAGGACAGCCCGGCTTCCA 1232
QY 66 -----LysCys-----Phe 68
Db 1233 TCAGTCCATCCCTTTCATTTGTCAAAGGGGATGCTCATCTCCAGGGCCTACTTC 1292
QY 69 MetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIlePro 88
Db 1293 GAGGCAGACATCAGTTAT-----1310
QY 89 GlyTyrProLysProAspMetAspAlaMetGluProSerTyrGluLeuIleHisThrGln 108
Db 1311 -----TCACAGGACCTTATTCGCTTTTAAACAGATGGATTCAGAGATATGAT 1361
QY 109 MetLysThrGlnGluTrp-----AspAsnSerLysSerIleLeuGlyVal 123
Db 1362 GTCAAGACCCAGGAAGTGGAGCTTCTCTTGGAAGACACAGTAAACTAATTGCAAGGTG 1421
QY 124 GlnCysGluValGlnLysGlnLysAlaPhe-----ValThrLeuGluArgPhe 140
Db 1422 CGCTGCCTCCCAAGTTGAGCTGGACCTCTGCCCCCAGCTCTCACCTGGCGTTTGTCT 1481
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThr-----157
Db 1482 TCTCAGCTCAAGAAGACATCTCTCAGTCTCACGCCAGATGTCCCAGAGGCAGACCTTTCT 1541
QY 158 -----LysLysPheAlaSerSerGlySerValPheGlyLys---GlyValLysPhe 173
Db 1542 GAAGTGGACCCCAAGCTCGTGTCTAATCTGATGCCCTTTTCAGAGAGCTGGAGTCAATTT 1601
QY 174 AlaLeu---LysAspGlyArgVal 180
Db 1602 GCCATAGCCAAAGGAGGCGCGCTG 1625

RESULT 10
US-08-162-809-15
; Sequence 15, Application US/08162809
; Patent No. 5457048
; GENERAL INFORMATION:
; APPLICANT: Pasquale, Elena B.
; APPLICANT: Sajjadi, Fereydoun G.
; TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States of America
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/162,809
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 9503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3254 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 32..2980
US-08-162-809-15

Alignment Scores:
Pred. No.: 9.49 Length: 3254
Score: 78.00 Matches: 58
Percent Similarity: 39.17% Conservative: 36
Best Local Similarity: 24.17% Mismatches: 82
Query Match: 7.24% Indels: 65
DB: 1 Gaps: 12

US-10-029-020-14_COPY_2400_2600 (1-201) x US-08-162-809-15 (1-3254)

QY 10 ProLeuThrLysLeuValHisMetGlyArgArgAspTyrAspValLeuAlaGlyArgTrp 29
Db 105 CCGGCAACGAAGTTAATCTGCTGGATTCAAACAATTC-----AAGGGAGCTGG 155
QY 30 Thr---SerProAspHisGluLeu--TrpLysHisLeuSer----- 41
Db 156 GCTGGATCTCCTACCCATCACATGGGTGGGAAGAGATTAGTGGTGTGATGAGCAATTATA 215
QY 42 -----SerSerAsnValMetProphe-----AsnLeuTyrMetP 53
Db 216 CTCCAATCAGAACTTACCAAGAGAGCAATGTTATGGATCACAGTCAAACAATTTGGCTGC 275
QY 53 heLysAsnAsnAsnProIleSerAsnSerGln-----AspIleLysCysPheM 69
Db 276 GAACAAACTGGATTCCACGCAATTCAGCGCAGAGATATATGTGGAGCTCAAGTTTACCT 335
QY 69 etThrAspValAsnSerTrpLeuLeuThrPheGly-----PheGlnLeuH 84
Db 336 TGAGGGACTGCAATAGTATCCCTCTAGTTCTGGGCACTTGCAAGAGACITTCATCTGT 395
QY 84 isAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluProSerTyrGluL 104
Db 396 AT-----TACATGGAATCCGATGATGACCATTTGGCAAAGTTCAGAGAGC 440
QY 104 euIleHisThrGlnMetLysThrGlnGluTrpAspAsnSer----- 117
Db 441 ACCAATTTACGAAGATTGACACCATTGGCGGTGATGAGAGCTTCACCCAGATGATCTTG 500
QY 118 --LysSerIleLeuGlyValGlnCysGluValGln----- 128
Db 501 GGGACCGGATTCCTCAAGCTGAATACCGAAGTCCGCGAGGTGGACCTGTTAGTAAGAAG 560
QY 129 -----LysGlnLeuLysAlaPheValThrLeuGluArgPheAspGlnL 143
Db 561 GCTTTTACTTGGCTTTCCAAGATGTAGTGTCATGTGTGCCTTAGTCTCGGTGCG-AGTG 619
QY 143 euTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPheAlaSerS 163
Db 620 TACTTCAAGAAGTGCCCTTTTCACTGTCTCAAGAACCTCGCCATGTTTCCAGATACAGTTCC 679
QY 163 erGly---SerValPheGlyLysGlyValLysPhe-----AlaLeuLysAspGlyA 179
Db 680 ATGGACTCCAGTCCCTGGTGGAGGTGCGGGGTCTTGTGTCAATCATCTCCAAGGAGGAA 739
QY 179 rgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAala 197
Db 740 GAGCCACCCAAGATGTACTG-----CAGCACGGAAGGAGAAATGGCTAGT 783

RESULT 11

US-09-738-946-11
; Sequence 11, Application US/09738946
; Patent No. 6579701
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: DROSOPHILA HOMOLOGUES OF GENES AND PROTEINS IMPLICATED IN CANCER
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: EX00-043C
; CURRENT APPLICATION NUMBER: US/09/738,946
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 60/170,832
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/170,838
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: 60/178,580
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/185,879
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/185,880
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/186,150
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/189,701
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 5706
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-738-946-11

Alignment Scores:
Pred. No.: 26.6 Length: 5706
Score: 77.50 Matches: 60
Percent Similarity: 38.46% Conservative: 30
Best Local Similarity: 25.64% Mismatches: 83
Query Match: 7.20% Indels: 61
DB: 4 Gaps: 12

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-738-946-11 (1-5706)

QY 1 IleGlyTyrHisGlyGlyLeu-TyrAspProLeuThrLysLeuValHisMetGlyArgAr 20
Db 3716 ATTGGCTACCATGGATTCTCTCAGTTGCTCTGTGTTAGAACAGTT-----CGAAG 3766
QY 20 gAspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLe 40
Db 3767 G-----GGATGTTATATAAATCTCTTAAGCCAAATTTGGACGCCTTT 3808
QY 40 uSerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnAsnProIleSe 60
Db 3809 ATCCTCAAGCAGTATATTTCCCGATTCCGACTTTATATTTGACTTTAAAAAATCGAACAC 3868
QY 60 rAsnSer-----GlnAspIleLysCysPheMet-ThrAspValAsnSerT 75
Db 3869 GCGAAAAACATAAAACTGCTGAACAGGCTGTAAAAAAGTTCATGCTCGAACATCGATGGAA 3928
QY 75 rpLeuLeuThrPheGly-----PheGlnLeuHisAsnValIleProGlyT 90
Db 3929 CTACTTTAAGCTTTTGAAGGGGAGCAAGTCACGGAAACATTCATCCATCAATAATCCCATTA 3988
QY 90 yrProLysProAspMet-----AspAlaMetGluProSerTyrGluLeuIleH 106
Db 3989 AAGCAACTCCGCCCATGTGGCGCTGCTCTAAGGTGATGCAATTACAGAGAGAGTACATC 4048
QY 106 isThrGlnMetLysThrGlnGlu-----T 114
Db 4049 CAACAATATTAAAGTTTCATTGGGAAGGAATTGTAGACCAAATGGTTTGGTTTAGAGAAAAGCT 4108
QY 114 rpAspAsnSerLysSerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAla- 133


```
Db 4109 GGACAGAG-----GAAGTTCTTCGACAACTACGCCAAG 4141
QY 134 -----PheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrI 148
Db 4142 GCCTAATTAATGCTATGCCATAGCCTTTGAAAAAAGGATAGTCTCAACATTCTACCA 4201
QY 148 leThrSerCysGlnGlnAlaProLysThrLysLysPheAlaSerSer---GlySerValP 167
Db 4202 TAACA-----CCTCACACGTTGCAATTTGTCAAAAAGCTGGGTTCTACGT 4246
QY 167 heGlyLysGlyValLysPheAlaLeuLysAspGlyArgValThrThrAspIleIle---S 186
Db 4247 TTGGCATTGGAAATAGAAAAATGTT-----CCGGATCAGTAACCTCCTCAATTTCTAATT 4300
QY 186 erValAlaAsnGluAspGlyArgValAlaAlaIle 198
Db 4301 CAGCAGCCTCGGAGTCTCTTGCTCGACGCGCCCAAGTT 4338

RESULT 12
US-09-252-991A-2392
; Sequence 2392, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 2392
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-2392

Alignment Scores:
Pred. No.: 2.82 Length: 1119
Score: 76.50 Matches: 35
Percent Similarity: 42.76% Conservative: 30
Best Local Similarity: 23.03% Mismatches: 47
Query Match: 7.10% Indels: 40
DB: 4 Gaps: 8

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-252-991A-2392 (1-1119)
QY 5 GlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArgAspTyrAspVal 24
Db 66 GGCAAGATCATCAGCCG---ACCCGCGTGCTCTAC---GGTGACCGCACCTACCCCTTC 119
QY 25 LeuAla-----GlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeuSer 41
Db 120 GTCGCCAAGTCGGTGGCGCTGGTCTCGCTCAACCCCGAGGTCAAGCCCGCCACGTGGTG 179
QY 42 SerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnAsnProIleSerAsn 61
Db 180 GTCGGCGGC-----CACTGCTTCATGCAGGAAGACCGCGCGATGCC 221
QY 62 SerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPhe 81
Db 222 GCCGAACGCGTAGCGGACTTTCTCTGCGCCCAATCCATGAACGCTCCTTCGCTGGGGAATC 281
QY 82 Gln-LeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro-- 100
Db 282 CGGCGCTCATTCATCGGTCCAGGAGGATGACCCACAGGACCGAGCCATGCGCCCCCT 341
QY 101 -----SerTyrGluLeuIle-----HisThrGlnMetLysThrGlnGluTrpAs 115
```

```
Db 342 CCTCTTCAGTGCCCTTCTCCTGCTTTCCGGGCATACCCAG-----GCCAGCGAATGGAA 395
QY 115 pAsnSerLysSerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheVa 135
Db 396 CGACAGCCAGGCCCGTG-----411
QY 135 lThrLeuGluArgPheAspGlnLeuTyrGlySer 146
Db 412 -----GACAAAGCTATTTCGGCGCG 429

RESULT 13
US-09-717-926-3
; Sequence 3, Application US/09717926
; Patent No. 6569657
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Cook, William J.
; TITLE OF INVENTION: 32140, A No. 6569657el Human Aldehyde
; TITLE OF INVENTION: Dehydrogenase
; FILE REFERENCE: 35800/205243
; CURRENT APPLICATION NUMBER: US/09/717,926
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/214,707
; PRIOR FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2772
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2772)
US-09-717-926-3

Alignment Scores:
Pred. No.: 11.7 Length: 2772
Score: 76.50 Matches: 31
Percent Similarity: 38.89% Conservative: 18
Best Local Similarity: 24.60% Mismatches: 48
Query Match: 7.10% Indels: 29
DB: 4 Gaps: 6

US-10-029-020-14_COPY_2400_2600 (1-201) x US-09-717-926-3 (1-2772)
QY 88 ProGlyTyrProLysProAspMetAspAlaMetGluProSerTyrGluLeuIleHisThr 107
Db 631 CCTCGTATACCCAGCCAGCAAGAGGGCA-----ACATATGAAGGTATCCAGAA 681
QY 108 GlnMetLysThrGln---GluTrpAsnSerLysSerIleLeu-----121
Db 682 AAGGAAAATGCTGAGATTTCTTGGGACCAGTCTGCCGAAAGTTTACATAACTGGATTCTGA 741
QY 122 -----GlyValGlnCysGluValGlnLysGlnLeuLysAlaPheVal 135
Db 742 GGTATGATAAAGTCCCTGGAGCTTGGACAGAGATAAATGGACAGATGGTCACTTTC--- 798
QY 136 ThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnAlaPro 155
Db 799 -----TATGGCTCGACATTACTGAATAGCTCTGTGCCTCCT 834
QY 156 -----LysThrLysLysPheAlaSerSerGlySerValPheGlyLysGlyVal 171
Db 835 GGAGAACCACTGGAAATTAAGAGTGGCAAGAGCCTGGTCTCTGTACCAAAAATGGACTT 894
QY 172 LysPheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGluAsp 191
Db 895 GTTCTTTTGTGTAACGATGGAAGACACTGACG---GTGAGAAATCTGCAGTTTGAAGAT 951
QY 192 GlyArgArgValAlaAla 197
Db 952 GGAAAAAATGATCCCTGCC 969
```


This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 2357.29 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGYHGGLYDPLTKLVHMGR.....TDIISVANEDGRRVAAILNH 201

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_13135@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1077	100.0	6246	29	AY413475	AY413475 Homo sapi
2	1071	99.4	5970	29	AY413476	AY413476 Pan trogl
3	1030	95.6	668	13	BQ563239	BQ563239 gi01b09.Y
4	951	88.3	734	13	BU233488	BU233488 603408464
5	896	83.2	637	14	CF171639	CF171639 B0845C10-
6	836	77.6	667	10	AW767437	AW767437 da69b01.Y
7	816	75.8	850	13	BU232300	BU232300 603341548
8	700	65.0	718	9	AL045768	AL045768 DKFZp434F
9	691	64.2	525	9	AL046228	AL046228 DKFZp434E
10	642.5	59.7	847	13	BU234988	BU234988 603408474
11	636.5	59.1	3038	11	AK037897	AK037897 Mus muscu
12	631.5	58.6	929	13	BU371366	BU371366 603598013
13	629	58.4	506	10	BE819779	BE819779 MRJ-BN036
14	618.5	57.4	621	14	CB578819	CB578819 AMGNNUC:N
15	566	52.6	776	14	CA777388	CA777388 ip19c06.X
16	566	52.6	5094	29	AY405420	AY405420 Homo sapi
17	552.5	51.3	616	13	BX506934	BX506934 DKFZp779I
18	527.5	49.0	521	14	CB720876	CB720876 AMGNNUC:N
19	511.5	47.5	885	13	BU120313	BU120313 603142826
20	502	46.6	555	12	BG732516	BG732516 333252 MA
21	498.5	46.3	728	14	CA344273	CA344273 674645 NC
22	498	46.2	720	13	BU346142	BU346142 603525314
23	484	44.9	428	14	CB794379	CB794379 AMGNNUC:N
24	470	43.6	5069	29	AY405421	AY405421 Pan trogl
25	466	43.3	625	10	BE865229	BE865229 UI-M-BH2.
26	456	42.3	748	9	AU169909	AU169909 AU169909
27	443.5	41.2	678	13	BQ572906	BQ572906 UI-M-FD0-
28	438.5	40.7	625	13	BX308610	BX308610 BX308610
29	438.5	40.7	732	14	CB519165	CB519165 UI-M-GH0-
30	436.5	40.5	624	13	BX308609	BX308609 BX308609
31	429.5	39.9	5087	29	AY405422	AY405422 Mus muscu
32	424	39.4	621	12	BI400254	BI400254 MI-P-AY1-
33	421.5	39.1	731	14	CF539420	CF539420 UI-M-GH0-
34	418.5	38.9	654	12	BM490220	BM490220 ppg2n.pk0
35	414	38.4	818	14	CD807778	CD807778 UI-M-GW0-
36	413.5	38.4	446	14	CB545337	CB545337 AMGNNUC:N
37	405.5	37.7	603	13	BX300081	BX300081 BX300081
38	397	36.9	473	14	CD736507	CD736507 4019565 1
39	388.5	36.1	564	9	AI820014	AI820014 wj60d07.x
40	386.5	35.9	650	10	BB576857	BB576857 BB576857
41	386.5	35.9	3760	11	AK031268	AK031268 Mus muscu
42	381.5	35.4	705	9	AA142988	AA142988 z148c01.r
43	380	35.3	232	14	T05481	T05481 EST03370 Fe
44	378.5	35.1	610	14	CB580731	CB580731 AMGNNUC:N
45	375.5	34.9	851	14	CB179955	CB179955 AGENCOURT

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AY413475
Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY413475
AY413475.1 GI:39769437
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6246)

6246 bp DNA linear GSS 12-DEC-2003

AUTHORS Clark,A.G., GJanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 6246)

AUTHORS Clark,A.G., GJanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..6246 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>6246 /locus_tag="HCM4903"

gene

ORIGIN

Alignment Scores: 6.19e-126 Length: 6246

Pred. No.: 1077.00 Matches: 201

Score: 100.00% Conservative: 0

Percent Similarity: 100.00% Mismatches: 0

Best Local Similarity: 100.00% Indels: 0

Query Match: 100.00% Gaps: 0

DB: 29

US-10-029-020-14_COPY_2400_2600 (1-201) x AY413475 (1-6246)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20

Db 5134 ATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTGTCTCCACATGGCGCGGA 5193

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40

Db 5194 GATTATGATGTGCTGGCGGACGCTGGACTAGCCAGACCCAGAGCTGTGGAAAGCACCTT 5253

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60

Db 5254 AGTAGCAGCAACGTCATGCCTTTTATCTCTATATGTTCAAAAACAACACCCCATCAGC 5313

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80

Db 5314 AACTCCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGTGCTCACCTTTGGA 5373

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100

Db 5374 TTCCAGCTACACAACGTCATCCCTGGTTATCCCAACCAGACATGGATGCCATGGAACCC 5433

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120

Db 5434 TCCTACGAGCTCATCCACACACACAGATGAAACCGCAGGAGTGGGACACAGCAAGTCTATC 5493

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140

Db 5494 CTCGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCCTTTGTCACTTAGAACGGTTT 5553

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160

Db 5554 GACCAGCTCTATGGTCCACATCACCAGCTGCCAGCAGGCTCCAAAGACCAAGAAGTTT 5613

QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180

Db 5614 GCATCCAGCGGCTCAGTCTTTTGGCAAGGGGTCAAGTTTGCTTTGAAGGATGGCCGAGTG 5673

QY 181 ThrThrAspIleSerValAlaAsnGluAspGlyArgValAlaAlaLeuAsn 200

Db 5674 ACCACAGACATCATCAGTGTGGCCATGAGGATGGCGAAGGGTGTGCTGCCATCTTGAC 5733

QY 201 His 201

Db 5734 CAT 5736

RESULT 2

AY413476 5970 bp DNA linear GSS 12-DEC-2003

LOCUS Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

DEFINITION AY413476 GI:39769438

ACCESSION AY413476

VERSION AY413476.1

KEYWORDS GSS.

SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes

REFERENCE 1 (bases 1 to 5970)

AUTHORS Clark,A.G., GJanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 5970)

AUTHORS Clark,A.G., GJanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..5970 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>5970 /locus_tag="HCM4903"

gene

ORIGIN

Alignment Scores: 3.4e-125 Length: 5970

Pred. No.: 1071.00 Matches: 200

Score: 99.50% Conservative: 0

Percent Similarity: 99.50% Mismatches: 1

Best Local Similarity: 99.50% Indels: 0

Query Match: 99.44% Gaps: 0

DB: 29

US-10-029-020-14_COPY_2400_2600 (1-201) x AY413476 (1-5970)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20

Db 4858 ATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCCACATGGCGCGGA 4917

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40

Db 4918 GATTATGATGTGCTGGCCGACGCTGGACTAGCCAGACCCAGAGCTGTGGAAGCACCTT 4977

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60

Db 4978 AGTAGCAGCAACGTCATGCCTTTTAACTCTCTATATGTTCAAAAACAACACCCCATCAGC 5037

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80

Db 5038 AACTCCCAGGACATCAAGTGCTTTCATGACAGATGTTAACAGCTGGCTGTACCTTTGGA 5097

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
|||||
Db 5098 TTCCAGCTACACAACGTGATCCCTGGTTATCCCAACACGACATGGATGCCATGGAACCC 5157

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
|||||
Db 5158 TCCTACGAGCTCATCCACACACAGATGAAACGCAGGAGTGGACAACAGCAAGTCTATC 5217

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
|||||
Db 5218 CTTGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTGTCACTTAGAACGGTTT 5277

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
|||||
Db 5278 GACCAGCTCTATGGTCCCAATCACCAGCTGCCAGCAGCTCAAAGACCAAGAGTTT 5337

QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
|||||
Db 5338 GCATCCAGCAGCTCAGTCTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTG 5397

QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaIleLeuAsn 200
|||||
Db 5398 ACCACAGACATCATCAGTGTGGCCATGAGGATGGCGAAGGTTGCTGCCATCTTGAAC 5457

QY 201 His 201
|||
Db 5458 CAT 5460

RESULT 3
BQ563239
LOCUS
DEFINITION
BQ563239
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Kachar,B.
EST analysis of gene expression in the mouse Organ of Corti at the
onset of hearing
Unpublished (2002)
Contact: Kachar,B.
Structural Cell Biology
National Institute of Deafness and other Communication Disorders
50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
Tel: 301-402-1599
Fax: 301-402-1765
Email: kachar@nidcd.nih.gov
Plate: 01 row: b column: 09
Seq primer: M13Rp1 reverse primer (ABI).

FEATURES
source
1..668
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="gi01b09"
/sex="male and female"
/dev_stage="Post natal day 5 to 13"
/clone_lib="Mouse Organ of Corti cDNA pBluescript"
/note="Organ: Organ of Corti; Vector: pBluescript; The organ of Corti (OC) was fine dissected from a total of 386 OC as follows: 102 samples from post-natal (p) day 5; 72 from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10; 14 from P12 and 24 from P13. After killing animals by cervical dislocation followed by decapitation, the bulla was removed and opened in Leibowitz medium. The bony capsule of the cochlea was shipped away, stria vascularis

and spiral ligament were removed and the sensory epithelium was carefully dissected out of the modiolus. Total RNA was extracted using the micro Fasttrack kit (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to manufacturer's instructions. Reverse transcription and library construction were carried out with the Uni-Zap XR vector kit (catalog # 237211, Stratagene) and Uni-Zap XR Gigapack III Gold Cloning kit (catalog # 237612), both from Stratagene (La Jolla, CA, USA), according to manufacturer's instructions. Briefly: 1.5 ug mRNA was reverse transcribed using a hybrid oligo(dT) linker-primer that contains an Xho I site. First strand synthesis was primed with the linker- primer and transcribed using Moloney murine leukemia virus reverse transcriptase (MMLV-RT) and 5-methyl dCTP. The second strand was synthesized with DNA polymerase and RNase H. Complementary DNA was blunt ended with Pfu DNA polymerase, ligated with EcoR I adapters in the presence of ligase and digested with Xho I. The cDNA was sequentially size fractionated over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden) and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA) columns to enrich for cDNAs greater than 400bp and 1000 bp, respectively. The cDNA was then directionally ligated to the Uni-ZAP XR vector, which had been predigested with EcoR I and Xho I. The phagemid was packaged with Gigapak III Gold and, upon titration on XL1 Blue MRF' cells, the yield of the phage library was estimated to be 11,100,000 recombinants. Stratagene's ExAssist Interference resistance helper phage (catalogue # 211203) was adopted to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CAGGAAACAGCTATGACC) and 25% strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have know function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."

ORIGIN

Alignment Scores:
Pred. No.: 2e-121 Length: 668
Score: 1030.00 Matches: 188
Percent Similarity: 99.00% Conservative: 11
Best Local Similarity: 93.53% Mismatches: 2
Query Match: 95.64% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x BQ563239 (1-668)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
:::|||||
Db 1 GTCGGCTACCACGCGGCCTCTATGATCCACTCACCAGCTTGTCACATGGGCCGACGG 60

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
|||||
Db 61 GATTATGATGTGCTGGCTGGACGCTGGACAAGCCACGACCATGAACTCTGGAACGCCTG 120

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
|||||
Db 121 AGTAGCACAGCATCGTGCCTTTTCATCTCTACATGTTTAAAGAACAAACCCCATCAGC 180

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80

methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

Db	181	AACTCTCAGGACATCAAGTGCTTCATGACAGATGTCAACAGCTGGCTCCTCACCTTTGGA	240		
QY	81	PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro	100		
Db	241	TTCCAGCTGCACAACGTGATACCTGGCTATCCCAAGCCAGACACAGATGCCATGGAACCC	300		
QY	101	SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle	120		
Db	301	TCCTACGAGCTCGTACACACACAGATGAAACTCAGGAATGGGACAAACAGCAAGTCTATC	360		
QY	121	LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe	140		
Db	361	CTCGGGGTACAGTGTGAAGTTCAGAAGCAACTCAAGGCTTCGTACCTTAGAACGCTTT	420		
QY	141	AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe	160		
Db	421	GACCAGCTCTACGGCTCGACCATCACCAGCTGCCAACAGGCCCTGAGACAAAGAAGTTT	480		
QY	161	AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal	180		
Db	481	GCCTCCAGTGGTTCATCTTTGGCAAGGGGTCAAGTTTGCTTGAAGATGGTCGAGTG	540		
QY	181	ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn	200		
Db	541	ACCACTGACATCATCAGTGTGGCCAATGAGGATGGCGGAGGATCGCAGCCATCTTGAAC	600		
QY	201	His	201		
Db	601	AAT	603		
RESULT 4	BU233488	734 bp	mRNA	linear	EST 26-NOV-2002
LOCUS	603408464F1	CSEQCHN24	Gallus gallus	cdna clone ChEST321p15 5', mRNA	
DEFINITION	sequence.				
ACCESSION	BU233488				
VERSION	BU233488.1	GI:25477142			
KEYWORDS	EST.				
SOURCE	Gallus gallus (chicken)				
ORGANISM	Gallus gallus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;				
	Phasianinae; Gallus.				
	1 (bases 1 to 734)				
	Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.				
	Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.				
	A Comprehensive Collection of Chicken cDNAs				
TITLE	Curr. Biol. 12 (22), 1965-1969 (2002)				
JOURNAL	22335534				
MEDLINE	12445392				
PUBMED	Contact: Simon Hubbard				
COMMENT	Department of Biomolecular Sciences				
	University of Manchester Institute of Science and Technology				
	(UMIST)				
	PO Box 88, Manchester, M60 1QD, UK				
	Tel: 01612008930				
	Fax: 01612360409				
	Email: Simon.Hubbard@umist.ac.uk.				
FEATURES	Location/Qualifiers				
source	1..734				
	/organism="Gallus gallus"				
	/mol_type="mRNA"				
	/strain="White Leghorn, Hisex"				
	/db_xref="taxon:9031"				
	/clone="ChEST321p15"				
	/dev_stage="22"				
	/lab_host="DH10B"				
	/clone_lib="CSEQCHN24"				
	/note="Organ: heads; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using				

ORIGIN

Alignment Scores:	3.31e-111	Length:	734
Pred. No.:	951.00	Matches:	174
Score:	94.92%	Conservative:	13
Percent Similarity:	88.32%	Mismatches:	10
Best Local Similarity:	88.32%	Indels:	1
Query Match:	88.30%	Gaps:	0
DB:	13		

US-10-029-020-14_COPY_2400_2600 (1-201) x BU233488 (1-734)

QY	1	IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg	20
Db	92	ATCGGCTACCATGGAGGACTGTATGATCCCCCTCACAAAGCTCATCCACATGGGACGGAGA	151
QY	21	AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu	40
Db	152	GACTATGATGTCTGGCAGGTGCGTGACAAAGTCCAGACCATGATATGTGAAGCACCTG	211
QY	41	SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer	60
Db	212	AGTAGCAATAACATCATGCCTTTCACACCTGTATATGTTCAAAACATAATAATCCCATCAGC	271
QY	61	AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly	80
Db	272	AACCTCTCAGGATATCAATGCTACATGACAGATGTCAACAGTTGGCTACTCAC-TTTGGG	330
QY	81	PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro	100
Db	331	TTCCAGCTACACAATGTCTATCCCTGGATACCCCAAGCCAGACCTGGATGCCATGGAGCCA	390
QY	101	SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle	120
Db	391	TCATATGAGCTTATCCACACGCAGATGAAACCCCAAGATGGACAAACAGCAAGTCAATT	450
QY	121	LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe	140
Db	451	TTAGGGGTCCATGTGAAGTGCAGAAAGCAGCTGAGGCCCTTTGTCACTCTTGAGCGCTTT	510
QY	141	AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe	160
Db	511	GAACAGATCTACAGTCCAGCATCGCGGTGCCAAGGTGTCAAGTTTGCCATGAAGGACGGCGTGT	570
QY	161	AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal	180
Db	571	GCCTCCGGAGGGTCTATCTTCGGCAAAAGGTGTCAAGTTTGCCATGAAGGACGGCGTGT	630
QY	181	ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAla	197
Db	631	GCACCGACATCATCAGTGTGCGCAATGAGGACGGCGGAGGATCGCAGCA	681

REFERENCE
AUTHORS
TITLE
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 637)
Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method

JOURNAL
MEDLINE
PUBMED
COMMENT
Genome Res. 11 (9), 1553-1558 (2001)
21429098
11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: B0845 row: C column: 10
Seq primer: M13 Reverse
High quality sequence stop: 637
POLYA=No.

FEATURES
source
1. .637
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="niaEST:B0845C10-5"
/db_xref="taxon:10090"
/clone="NIA:B0845C10 IMAGE:30471777"
/dev_stage="Newborn Kidney"
/lab_host="DH10B"
/clone_lib="NIA Mouse Newborn Kidney cDNA Library (Long
1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
In brief, double-stranded cDNAs were synthesized with an
Oligo(dT) primer [Invitrogen:
5'-pGACTAGTTCTAGATCGCGAGCGGCCCTTTT-3'] from
26 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker lI-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 3.0 kb. The library was
constructed by Yulan Piao."

ORIGIN

Alignment Scores:
Pred. No.: 3.11e-104 Length: 637
Score: 896.00 Matches: 162
Percent Similarity: 98.84% Conservatives: 8
Best Local Similarity: 94.19% Mismatches: 2
Query Match: 83.19% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x CF171639 (1-637)

QY 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
120 ATCGGCTACCAACGCGCGCTCTATGATCCACTCACCAAGCTTGTCACATGGCCGACGG 179
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
180 GATTATGATGTGCTGGCTGGACGCTGGACACAGCCAGACCATGAACCTCTGGAAACGCCTG 239
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
240 AGTAGCAACAGCATCGTGCCTTTTCATCTCTACATGTTTAAAGAACAAACCCCATCAGC 299

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 300 AACTCTCAGGACATCAAGTGTTCATGACAGATGTCAACAGCTGGCTCCTCACCTTTGGA 359
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 360 TTCCAGCTGCACAACAGTGATACCTGGCTATCCCAAGCCAGACACAGATGCCATGGAACCC 419
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 420 TCCTACGAGCTCGTACACACACAGATGAAAACCTCAGGAATGGACAACAGCAAGTCTATC 479
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 480 CTCGGGTACAGTGTGAAGTTTCAGAAAGCAACTCAAGGCTTTCGTTACCTTAGAACGCTTT 539
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 540 GACCAGCTCTACGCTCGACCATCACCAGCTGCCAACAGGCCCTGAGACAAAGAAAGTTT 599
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLys 172
Db 600 GCCTCCAGTGGTTCATCTTTGGCAAGGGGGTCAAG 635

RESULT 6
AW767437
LOCUS
DEFINITION
AW767437 667 bp mRNA linear EST 16-FEB-2001
da69b01.y1 Harland stage 19-23 Xenopus laevis cDNA clone
IMAGE:3200137 5' similar to TR:O70465 O70465 DOC4. ;, mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AW767437.1 GI:7699442
EST.
Xenopus laevis (African clawed frog)
Xenopus laevis

REFERENCE
AUTHORS
Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L.,
Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
Waterston,R. and Wilson,R.
WashU Xenopus EST project, 1999
Unpublished (1999)
Contact: Sandy Clifton, Ph.D.
WashU Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

TITLE
JOURNAL
COMMENT
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 494.

FEATURES
source
1. .667
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:3200137"
/tissue_type="neurula"
/dev_stage="stage 19-23"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Harland stage 19-23"
/note="Vector: pCS107 (custom); Site 1: NotI; Site 2:
SalI; cDNA made by oligo-dT priming. Library constructed
Library constructed by R. Harland, PhD. (University of California,
Berkeley)
Email: est@watson.wustl.edu

Library constructed by R. Harland, PhD. (University of California,
Berkeley)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -40RP from Gibco
High quality sequence stop: 494.

Location/Qualifiers
1. .667
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:3200137"
/tissue_type="neurula"
/dev_stage="stage 19-23"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Harland stage 19-23"
/note="Vector: pCS107 (custom); Site 1: NotI; Site 2:
SalI; cDNA made by oligo-dT priming. Library constructed

by Dr. Francesca Mariani in the laboratory of R. Harland, Ph.D. (University of California, Berkeley). References: XBF-2 is a transcriptional repressor that converts ectoderm into neural tissue. Mariani, FV. Harland, RM., Development. 1998 Dec;125(24):5019-31. PMID: 9811586; UI: 99030283; Use of large-scale expression cloning screens in the xenopus laevis tadpole to identify gene function. Grammer TC, Liu KJ, Mariani FV, Harland RM, Dev Biol. 2000 Dec 15;228(2):197-210. PMID: 11112324; UI: 20564075; Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores: 1.73e-96 Length: 667
Pred. No.: 836.00 Matches: 149
Score: 93.41% Conservative: 21
Percent Similarity: 81.87% Mismatches: 12
Best Local Similarity: 77.62% Indels: 0
Query Match: 10 Gaps: 0
DB:

US-10-029-020-14_COPY_2400_2600 (1-201) x AW767437 (1-667)

QY 20 ArgAspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHis 39
Db 2 CGGGAATATGATGCTTAGCTGGCGATGGACGAGTCCAGACCATGATATATGGAAGCGC 61
QY 40 LeuSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIle 59
Db 62 TTGAGTACCACCAATGTAATGCCGTTCAACCTCTACATGTTTAAAGAACAAACCCAGTT 121
QY 60 SerAsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuThrPhe 79
Db 122 AGCAATGCTCAAGACACCAATGTTACATGACAGATGTGAACAGTTGGCTGCTGACTTTC 181
QY 80 GlyPheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGlu 99
Db 182 GGCTTTTCAGCTGCACAATGTGATCCCCGGATACCCGAAACAGATATAGATGCCATGGAG 241
QY 100 ProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSer 119
Db 242 CCGTCCTATGAACACTCAACACGACAGATGAAGACTCAGGAGTGGGACACAGTAAGTCT 301
QY 120 IleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArg 139
Db 302 ATTNTGGAGTTTCAGTGTGAGGTGCAGAAACAACTGAAGGCATTCGTCACCTTGGAGCGC 361
QY 140 PheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys 159
Db 362 TTCNGTCAGATTACGGCTCCACTATGGCCGGCTGTCACCTTGAATCAGAGACGAGGAAG 421
QY 160 PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArg 179
Db 422 TTGCTTCTCTTGATCCATCTTTGGAAGGGCGTCAAGTTGCCATGAAAGACGGAAGG 481
QY 180 ValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeu 199
Db 482 ATAACCACCGATATCATCAGCGTCGCTAACGAGGATGGCAGGAGGATTGCTGCCATNGTG 541
QY 200 AsnHis 201
Db 542 AACAAAC 547
RESULT 7
BU232300
LOCUS 850 bp mRNA linear EST 26-NOV-2002
DEFINITION 603341548F1 CSEQCHN24 Gallus gallus cdna clone ChEST255011 5', mRNA
sequence.
ACCESSION BU232300
VERSION BU232300.1 GI:25474859
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

1 (bases 1 to 850)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

22335534
12445392

Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. .850

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hisex"

/db_xref="taxon:9031"

/clone="ChEST255011"

/dev_stage="22"

/lab_host="DH10B"

/clone_lib="CSEQCHN24"

/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand reaction, double-stranded cDNA
Following this first strand reaction, digested with
was bluntended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores: 9.43e-94 Length: 850
Pred. No.: 816.00 Matches: 148
Score: 94.15% Conservative: 13
Percent Similarity: 86.55% Mismatches: 10
Best Local Similarity: 75.77% Indels: 0
Query Match: 13 Gaps: 0
DB:

US-10-029-020-14_COPY_2400_2600 (1-201) x BU232300 (1-850)

QY 31 SerProAspHisGluLeuTrpLysHisLeuSerSerSerAsnValMetProPheAsnLeu 50
Db 7 TCGAGAGACCATGATGATGTGGAGCACCTGAGTAGCAATAACATCATGCTTTCACACCTG 66

QY 51 TyrMetPheLysAsnAsnAsnProIleSerAsnSerGlnAspIleLysCysPheMetThr 70
Db 67 TATATGTTCAAAAACAATAATCCCATCAGCAACTCTCAGGATATCAAATGCTACATGACA 126

QY 71 AspValAsnSerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIleProGlyTyr 90
Db 127 GATGTCACACAGTTGGCTACTCACTTTTGGGTTCCAGCTACACAATGTCTCCCTGGATAC 186

QY 91 ProLysProAspMetAspAlaMetGluProSerTyrGluLeuIleHisThrGlnMetLys 110
Db 187 CCCAAGCCAGACCTGGATGCCATGGAGCCATCATATGAGTTATCCACAGCAGATGAAA 246

QY 111 ThrGlnGluTrpAspAsnSerLysSerIleLeuGlyValGlnCysGluValGlnLysGln 130
Db 247 ACCCAAGAATGGACACACAGCAAGTCAATTTTAGGGGTCCCAATGTGAAGTGCAGAAGCAG 306

QY	131	LeuLysAlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThrIleThrSer	150
Db	307	CTGAAGGCCTTTGTCACTCTTGAGCGCTTTGAACAGATCTACAGCTCCAGCATCGCCGG	366
QY	151	CysGlnGlnAlaProLysThrLysLysPheAlaSerSerGlySerValPheGlyLysGly	170
Db	367	TGCCAACAGGTCAAGAAGAACAAAGAACTTCGCCTCCGAGGGTCTATCTTCGGCAAAGT	426
QY	171	VallysPheAlaLeuLysAspGlyArgValThrThrAspIleIleSerValAlaAsnGlu	190
Db	427	GTCAGTTTGCCATGAAGACGGCGTGTTCACCCGACATCATCAGTGTGGCCATGAG	486
QY	191	AspGlyArgArgValAlaAlaIleLeuAsnHis	201
Db	487	GACGGCGGAGGATCGCAGCATCTTGAACAAT	519
RESULT	8		
AL045768			
LOCUS	AL045768		
DEFINITION	DKFZp434F206_r1 434 (synonym: htes3) Homo sapiens cDNA clone	718 bp	mRNA linear EST 04-SEP-2003
ACCESSION	AL045768		
VERSION	AL045768.1	GI:5433880	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 718)		
TITLE	Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.		
JOURNAL	EST (Koehrer, et al.)		
COMMENT	Unpublished (1999)		
	Contact: MIPS		
	MIPS		
	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany		
	This is the 5' sequence of the clone insert		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;		
	sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.		
	s1 sequence also available.		
	This clone (DKFZp434F206) is available at the RZPD in Berlin.		
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.		
FEATURES	Location/Qualifiers		
source	1..718		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DKFZp434F206"		
	/tissue_type="testis"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/clone_lib="434 (synonym: htes3)"		
	/note="Vector: pSport1; Site_1: NotI; Site_2: Sali"		
ORIGIN			
Alignment Scores:			
Pred. No.:	5.89e-79	Length:	718
Score:	700.00	Matches:	134
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	65.00%	Indels:	0
DB:	9	Gaps:	0
US-10-029-020-14_COPY_2400_2600 (1-201) x AL045768 (1-718)			
QY	68	PheMetThrAspValAsnSerTrpLeuLeuThrPheGlyPheGlnLeuHisAsnValIle	87
Db	3	TTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTGATC	62
QY	88	ProGlyTyrProLysProAspMetAspAlaMetGluProSerTyrGluLeuIleHisThr	107

Db	63	CCTGGTTATCCAAACCAGACATGGATGCCATGGAAACCCCTCCTACGAGCTTATCCACA	122
QY	108	GlnMetLysThrGlnGluTrpAspAsnSerLysSerIleLeuGlyValGlnCysGluVal	127
Db	123	CAGATGAAACCGCAGGAGTGGGACACAGCAAGTCTATCCTCGGGTACAGTGTGAAGTA	182
QY	128	GlnLysGlnLeuLysAlaPheValThrLeuGluArgPheAspGlnLeuTyrGlySerThr	147
Db	183	CAGAAGCAGCTCAAGGCCCTTTGTACCTTAGAACGGTTTGACCATCTATGGCTCCACA	242
QY	148	IleThrSerCysGlnGlnAlaProLysThrLysLysPheAlaSerSerGlySerValPhe	167
Db	243	ATCACAGCTGCCAGCGAGCTCCAAAGACCAAGAAAGTTTGCATCCAGCGGTCTTT	302
QY	168	GlyLysGlyVallysPheAlaLeuLysAspGlyArgValThrThrAspIleIleSerVal	187
Db	303	GGCAAGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTGACCAACATCATCATGTG	362
QY	188	AlaAsnGluAspGlyArgValAlaAlaIleLeuAsnHis	201
Db	363	GCCAATGAGGATGGCGAAGGTTTGCTGCCATCTTGAACCAT	404
RESULT	9		
AL046228			
LOCUS	AL046228		
DEFINITION	DKFZp434E177_r1 434 (synonym: htes3) Homo sapiens cDNA clone	525 bp	mRNA linear EST 04-SEP-2003
ACCESSION	AL046228		
VERSION	AL046228.1	GI:5434312	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 525)		
TITLE	Koehrer, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.		
JOURNAL	EST (Koehrer, et al.)		
COMMENT	Unpublished (1999)		
	Contact: MIPS		
	MIPS		
	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany		
	This is the 5' sequence of the clone insert		
	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;		
	sequenced by BMFZ (Biomedical Research Center at the Charite, Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.		
	No s1 sequence available.		
	This clone (DKFZp434E177) is available at the RZPD in Berlin.		
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.		
FEATURES	Location/Qualifiers		
source	1..525		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="DKFZp434E177"		
	/tissue_type="testis"		
	/dev_stage="adult"		
	/lab_host="DH10B"		
	/clone_lib="434 (synonym: htes3)"		
	/note="Vector: pSport1; Site_1: NotI; Site_2: Sali"		
ORIGIN			
Alignment Scores:			
Pred. No.:	5.15e-78	Length:	525
Score:	691.00	Matches:	133
Percent Similarity:	99.25%	Conservative:	0
Best Local Similarity:	99.25%	Mismatches:	1
Query Match:	64.16%	Indels:	0
DB:	9	Gaps:	0

reannealing hybridization was used."

QY 119 SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu 138
Db 2220 ATGATCCTGGGCATTGAGCTCCAGAACAACTAAGGAATTTTCATTCCTGGAT 2279
QY 139 ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnAlaProLysThrLys 158
Db 2280 CAGCTTCCTATGACTCCTCAGTACATGAGGGAGGTGTCTTGAAGGAGGAAACAGCCG 2339
QY 159 LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly 178
Db 2340 AGGTTTGCTGCTGCTCCCTTCTGCTTGGCAAAGGTATCAAATTTGCCATCAAGGAGGC 2399
QY 179 ArgValThrThrAspIleLeuSerValAlaAsnGluAspGlyArgValAlaAlaLalle 198
Db 2400 ATAGTGACAGCTGATATTATAGGAGTAGCCAATGAAGATAGCAGCGCTTGTGTCATT 2459
QY 199 LeuAsnHis 201
Db 2460 CTCACAAT 2468

RESULT 12
BU371366
LOCUS 603598013F1 CSEQCHN73 Gallus gallus cdna clone CHEST56814 5', mRNA
DEFINITION sequence.
ACCESSION BU371366
VERSION BU371366.1 GI:25879367
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
REFERENCE Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
AUTHORS Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .929
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHEST56814"
/sex="Female"
/tissue_type="not cerebrum or cerebellum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQCHN73"
/note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was bluntended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer

ORIGIN

Alignment Scores:

Pred. NO.: 5.65e-70 Length: 929
Score: 631.50 Matches: 111
Percent Similarity: 76.96% Conservatives: 36
Best Local Similarity: 58.12% Mismatches: 43
Query Match: 58.64% Indels: 1
DB: 13 Gaps: 1

US-10-029-020-14_COPY_2400_2600 (1-201) x BU371366 (1-929)

QY

1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 84 ATTGGTTTTCATGGAGGGCTCTATGATTCTCTCACCACCAATTGGTTTCATCTGGGTCAAGCG 143

QY

21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
Db 144 GATTATGATGTTATTGCTGCTGGTGGTGGACACACCAACCATCATATATGGAACACCTG 203

QY

41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 204 AATGCTGTC--CCACAACCAATCAATCTCTACTCATTTTGAATAATACTACCCAGTTGGC 260

QY

61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 261 AGGATCCAAGATGTTGCTAAGTATACACAGACATTTGGAAGTTGGCTAGAGCTATTGGT 320

QY

81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 321 TTCCAGTTGGACAATGTACTACCTGGGTTCACAAACCAAGAAATCGAAGCTTTGGAGACA 380

QY

101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 381 ACATATGAACCTTCTACAGCTTCAACCAAAACCCAGGAGTGGATCCTCGAAAGACTATC 440

QY

121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 441 CTTGGTATTTCAGTGTGAGCTACAGAAGCAACTCCGAAACTTTATATCCTTGGATCAACTT 500

QY

141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 501 CCAATGACCCCGAGGTATAGTATGATGGCAAGTCTATGAGGGAGTGAAGCAACCGAGTTT 560

QY

161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 561 GCAGCTATTTCCTTCAGTATTGGAAAAGGCAATCAAAATTTGCTATCAAGGATGGCATCGTC 620

QY

181 ThrThrAspIleIleSerValAlaAsnGluAsp 191
Db 621 ACAGCGGACATTATTGTTGTTGCTAATGAGGAC 653

RESULT 13

BE819779/c 506 bp mRNA linear EST 21-SEP-2000

LOCUS

MR3-BN0363-150800-006-c07 BN0363 Homo sapiens cdna, mRNA sequence.

DEFINITION

BE819779

ACCESSION

BE819779

VERSION

BE819779.1 GI:10252013

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

1 ileGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
591 ATTGATTTTCATGGTGGCTGTATGACCCACTCACCAATTAATCCACTTTGGAGAAAGA 591
21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTyrLysHisLeu 40
590 GATTATGACATTTTGGCAGGACGGTGGACACACCTGACATAGAAATCTGGAAGAAT 531
41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
530 ---GGGAAGGACCCAGCTCTTTTAACTTTGTACATGTTTAGGAATAACAACCCCTGCAAGC 474
61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTyrLeuLeuThrPheGly 80
473 AAAATCCATGACGTGAAGATTACATCACAGATGTTAACAGCTGGTGGTGACATTGGT 414
81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
413 TTCCATCTGCACAATGCTATTCTCTGGATTCCCTGTTCCCAAAATTTGATTAAACAGAACCT 354
101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
353 TCTTACGAACCTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 306
121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
305 TTCGAGTCCAGCAGCAAGTGGCGCGCAGGCCCAAGGCCCTTCTCTGCTGGGGAAGATG 246
141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
245 GCCGAGGTGCAG-----GTGAGCCGCGCGCGCGCGCGCGCGCGCGCTCTGG 198
160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
197 CTGTGGTTCGCCACGGTCAAGTCGCTGATCGGCAAGGGCGTCAATGCTGGCCGTCAGCCAG 138
178 GlyArgValThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgValAlaAla 197
137 GGCCGCGTGCAGACCAACAGTGTCTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGGCC 78
198 IleLeuAsnHis 201
77 GTGCTCAACAAC 66

Search completed: August 14, 2004, 18:04:53
Job time : 2373.29 secs

163 SerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgValThrThr 182
472 GTCCCTTCTGCTTGGCAAGGTATAAAATTTGCCATCAAGGAGGCGATAGTACAGCT 531
183 AspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIleLeuAsnHis 201
532 GATATTATAGGAGTAGCCCAATGAAGATAGCAGGCGTCTGTGCTGCCATTCTCAACAAT 588

RESULT 15
CA777388/c 776 bp mRNA linear EST 03-DEC-2002
LOCUS ip19c06.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6217715 3'
DEFINITION similar to TR:Q9WTS6 Q9WTS6 TEN-M3. ; mRNA sequence.

ACCESSION CA777388
VERSION CA777388.1 GI:26015263

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 776)

AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.

TITLE Endocrine Pancreas Consortium

JOURNAL Unpublished (2000)

COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)

Seq primer: -40UP from Gibco

High quality sequence stop: 439.

FEATURES

source

1..776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6217715"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@ingate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

Alignment Scores:

Pred. No.: 1.12e-61 Length: 776
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 14 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x CA777388 (1-776)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - nucleic search, using frame_plus_p2n model
Run on: August 13, 2004, 17:03:51 ; Search time 340.95 Seconds
(without alignments)
2504.436 Million cell updates/sec
Title: US-10-029-020-14_COPY_2400_2600
Perfect score: 1077
Sequence: 1 IGYHGGLYDPLTKLVHMGR.....TDIISVANEDGRRVAAILNH 201

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029020@cgn 1 1 1868 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	1077	100.0	8354 6 ABS52100 Human TEN
2	1077	100.0	8645 6 ABS78652 Human CDN
3	1063	98.7	8438 6 ABN85378 Human NOV
4	642.5	59.7	13202 4 AAK51828 Human pol
5	630.5	58.5	12879 6 ABK92230 Prostate
6	566	52.6	2157 4 AAH14096 Human CDN
7	566	52.6	3270 4 AAH14671 Human CDN
8	566	52.6	3614 4 AAH14183 Human CDN

9	566	52.6	8473	6	ABQ82345	Abq82345 Human NOV
10	566	52.6	8487	6	ABQ82346	Abq82346 Human NOV
11	566	52.6	8645	6	ABQ82344	Abq82344 Human NOV
12	566	52.6	8675	6	ABQ82343	Abq82343 Human NOV
13	556	51.6	791	6	ABK34316	Abk34316 Human CDN
14	529.5	49.2	9058	7	ACC72051	Acc72051 BCU0205A
15	529.5	49.2	9695	7	ACC72052	Acc72052 BCU0205B
16	529.5	49.2	9729	5	AAS14089	Aas14089 Human FCT
17	529.5	49.2	9729	9	ADB32028	Adb32028 Human FCT
18	529.5	49.2	9826	5	AAS14085	Aas14085 Human FCT
19	529.5	49.2	9826	9	ADB32023	Adb32023 Human FCT
20	497.5	46.2	3312	4	AAK52812	Aak52812 Human pol
21	200.5	18.6	10242	4	ABL29075	Ab129075 Drosophil
22	200.5	18.6	17131	4	ABL29074	Ab129074 Drosophil
23	194	18.0	431	6	ABL65724	Ab165724 Lung canc
24	188.5	17.5	1973	4	AAI14695	Aai14695 Probe #46
25	188.5	17.5	1973	4	ABA56427	Aba56427 Human foe
26	188.5	17.5	1973	4	AAI36063	Aai36063 Probe #47
27	188.5	17.5	1973	4	ABA45900	Aba45900 Human bre
28	188.5	17.5	1973	4	ABA26060	Aba26060 Probe #45
29	188.5	17.5	1973	4	AAK30099	Aak30099 Human bon
30	188.5	17.5	1973	4	AAK04592	Aak04592 Human bra
31	188.5	17.5	1973	4	ABS29749	Abs29749 Human liv
32	188.5	17.5	1973	5	AAI04501	Aai04501 Probe #44
33	188.5	17.5	1973	6	ABS04679	Abs04679 Human gen
34	174	16.2	728	5	AAS68860	Aas68860 DNA encod
35	164.5	15.3	3910	4	ABL04849	Ab104849 Drosophil
36	164.5	15.3	6318	4	ABL04848	Ab104848 Drosophil
37	131.5	12.2	399	6	AAS20981	Aas20981 DNA seque
38	128	11.9	464	8	ACH40709	Ach40709 Human foe
39	87.5	8.1	2836	7	ABX70662	Abx70662 Human CDN
40	86	8.0	1630	4	ABL03289	Ab103289 Drosophil
41	84.5	7.8	1222	4	AAD03763	Aad03763 Human Ig4
42	84.5	7.8	1527	7	ABX06099	Abx06099 S. pneumo
43	84.5	7.8	20986	2	AAV52187	Aav52187 Streptoco
44	84.5	7.8	110000	7	ABS56454_03	Continuation (4 of
45	84	7.8	2419	3	AAZ91917	Aaz91917 Murine ma

ALIGNMENTS

RESULT 1	
ABS52100	ABS52100 standard; DNA; 8354 BP.
ID	ABS52100 standard; DNA; 8354 BP.
XX	
AC	ABS52100;
XX	
DT	05-NOV-2002 (first entry)
XX	
DE	Human TEN-M4-like gene.
XX	
KW	Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW	cell signal processing; metabolic pathway modulation; metabolic disorder;
KW	obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW	Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW	haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW	memory defect; infertility; congenital heart defect; hair growth;
KW	pigmentation disorder; endocrine disorder; respiratory disease; health;
KW	gastro-intestinal disease; reproductive; neurological disease;
KW	bone marrow transplantation; endocrine disease; allergy; inflammation;
KW	nephrological disorder; urinary system disorder; age-related disorder;
KW	neuropsychiatric disorder; EGF-related protein; SCUBEL; TEN-M4;
KW	adipocyte complement-related C1q tumour necrosis factor; out at first;
KW	beta adrenergic receptor kinase; BphA6/ehk-2; glucose transporter;
KW	type 1a membrane sushi-containing domain; butyrophilin;
KW	type 1a membrane-sushi domain containing; SNP; gene; ds;
KW	single nucleotide polymorphism.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	variation replace(117,G)
FT	/*tag= a

```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
XX 25-JUL-2002.
XX
XX 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
XX
XX 20-DEC-2000; 2000US-0257314P.
XX
XX 02-MAY-2001; 2001US-0288153P.
XX
XX 29-MAY-2001; 2001US-0294075P.
XX
XX 24-JUL-2001; 2001US-0307506P.
XX
XX 10-AUG-2001; 2001US-0311590P.
XX
XX 10-AUG-2001; 2001US-0311613P.
XX
XX 29-AUG-2001; 2001US-0315617P.
XX
XX 14-SEP-2001; 2001US-0322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
XX Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
XX Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX WPI; 2002-590744/63.
XX P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
XX atherosclerosis, metabolic disorders, diabetes, obesity, infectious
XX disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
XX cancer.
XX
XX Claim 8; Page 50-52; 318pp; English.
XX
XX The present invention relates to new NOVX polypeptides. The invention is
XX useful for treating or preventing a NOVX-associated disorder such as
XX cardiomyopathy or atherosclerosis, where the disorder is related to cell
XX signal processing and metabolic pathway modulation in a subject,
XX preferably human. The invention is also useful for treating metabolic
XX disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
XX disorders, haematopoietic disorders and various cancers. The molecules of
XX the invention are also useful for treating or preventing cirrhosis,
XX pancreatitis, learning and memory defects, infertility, congenital heart
XX defects, acne, hair growth, pigmentation disorders, endocrine disorders,
XX respiratory disease, gastro-intestinal diseases, reproductive health,
XX neurological diseases, bone marrow transplantation, endocrine diseases,
XX allergy and inflammation, nephrological disorders, urinary system
XX disorders, neuropsychiatric disorders and age-related disorders. The
XX present nucleic acid sequence represents a NOVX gene. This sequence
XX encodes a NOVX protein of the invention
XX
XX Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. NO.: 6.8e-122 Length: 8354
XX Score: 1077.00 Matches: 201
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-10-029-020-14_COPY_2400_2600 (1-201) x ABS52100 (1-8354)
XX
XX 1 IleGlyTyrHisGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
XX
XX 7232 ATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAGCTGTCCACATGGGCCGCGCA 7291
XX
XX
```

```
QY 21 AspTyrAspValLeuAlaGlyArgTirThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7292 GATTATGATGTGCTGGCCGGACGCTGGACTAGCCACAGACCACGAGCTGTGGAAGCACCTT 7351
QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7352 AGTAGCAGCAACGTCATGCCTTTTAAATCTCTATATGTTCAAAAACAACACCCCATCAGC 7411
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTirPLeuLeuThrPheGly 80
Db 7412 AACTCCCAGGACATCAAGTGCTTTCATGACAGATGTTAAACAGCTGGCTGCCTTTTGA 7471
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7472 TTCCAGCTACACAAACGTCATCCCTGGTTATCCCAAACACAGACATGGATGCCATGGAACCC 7531
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7532 TCCTACGAGCTCATCCACACACAGATGAAACGCGAGAGTGGGACACAGCAAGTCTATC 7591
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLysLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7592 CTCGGGTACAGTGTGAAGTACAGAACGACGCTCAAGGCCTTTGTACACCTTAGAACGGTTT 7651
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 7652 GACCAGCTCTATGGCTCCACAATCACCAGCTGCCAGCAGCTCCAAAGACCAAGAAGTTT 7711
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 7712 GCATCCAGCGGCTCAGTCTTTGGCAAGGGGGTCAAGTTTGCTTGAAGGATGGCCGAGTG 7771
QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
Db 7772 ACCACAGACATCATCAGTGTGGCCAATGAGGATGGCGAAGGGTGTGCTGCCATCTTGAAC 7831
QY 201 His 201
Db 7832 CAT 7834
RESULT 2
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
AC ABS78652;
XX
XX 16-DEC-2002 (first entry)
XX Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
XX Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
XX polycythaemia vera; primary thrombocytopaenia; developmental disorder;
XX renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
XX neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
XX reproductive disorder; infertility; autoimmune disorder; gout; allergy;
XX inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
XX autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
XX diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
XX multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
XX rheumatoid arthritis.
XX
XX Homo sapiens.
XX WO200272830-A2.
XX
XX 19-SEP-2002.
XX
XX 08-FEB-2002; 2002WO-US0003715.
XX
XX 09-FEB-2001; 2001US-0268111P.
XX
XX
```

PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
PR 09-MAR-2001; 2001US-0274552P.
XX (INCY-) INCYTE GENOMICS INC.
PA
XX
PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
DR WPI; 2002-723356/78.
DR P-PSDB; ABG97359.
XX
PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX
PS Claim 5; Page 175-178; 181pp; English.
XX
CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 7.14e-122 Length: 8645
Score: 1077.00 Matches: 201
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_2400_2600 (1-201) x ABS78652 (1-8645)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArg 20
Db ATAGGCTACCATGGTGGCCTCTATGATCCACTCACCAAGCTTGTCACATGGCCGCGA 7340

QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db GATTATGATGTGTCGCGGACGCTGGACTAGCCAGACCACGAGCTGTGGAAGCACCTT 7400

QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db AGTAGCAGCAACGTCATGCCTTTTAACTCTATATGTTCAAAACAAACACCCCATCAGC 7460

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db AACTCCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGGCTGCCTACCTTGA 7520

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db TTCCAGCTACACAACGTCATCCCTGGTTATCCAAACACAGATGGATGCCATGGAACCC 7580

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db TCCTACGAGCTCATCCACACACAGATGAAACGCGAGGAGTGGACAAACAGCAAGTCTATC 7640

QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db CTCGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCCCTTTGTCCCTTAGAACGGTTT 7700

QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db GACCAGCTCTATGGCTCCACAATCACCAGCTGCCAGGCTCCAAAGACCAAGAAGTTT 7760

QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db GCATCCAGCGGCTCAGTCTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGGATGGCCGAGTG 7820

QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
Db ACCACAGACATCATCAGTGTGGCCAATGAGGATGGGCGAAGGTTGCTGCCATCTTGAAC 7880

QY 201 His 201
Db CAT 7883

RESULT 3
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
AC ABN85378;
XX 21-OCT-2002 (first entry)
XX Human NOV1, TEN-M4 like protein, coding sequence.
XX Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IgA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX Homo sapiens.
OS
XX
XX FH Key Location/Qualifiers
FT CDS 4..8395
FT /*tag= a
FT /trans_except= (pos: 1138..1147,aa:Met)
FT /product= "NOV1 protein"
XX
PN WO200255704-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-US0000554.
XX
PR 09-JAN-2001; 2001US-0260417P.
PR 10-JAN-2001; 2001US-0260831P.
PR 28-FEB-2001; 2001US-0272338P.
PR 09-MAR-2001; 2001US-0274876P.
PR 18-APR-2001; 2001US-0284704P.
XX

QY 79 pheGlypheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMet 98
Db 7364 TTTGGTTTCCAAATACACAATGTACTACCTGGATTCCCAACCTGAATAGAAATTTA 7423
QY 99 GluProSerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLys 118
Db 7424 GAATTAACCTACGAGCTTCTACGGCTTCAGACAAAAAACTCAAGAGTGGGATCCTGGAAAG 7483
QY 119 SerIleLeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGlu 138
Db 7484 ACTATCCTGGGCAATTCAGTGTGAACCTCCAGAAACAGCTCAGGAATTTCAATTCCTTGGAC 7543
QY 139 ArgPheAspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLys 158
Db 7544 CAACTACCTATGACTCCCCGATACATGATGGACGGTGCCTTGAAGGAGGGAAGCAACCA 7603
QY 159 LysPheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGly 178
Db 7604 AGGTTTGCTGCTGTCCCTTCTGTTTGGGAAAGGTATATAAAATTTGCCATCAAGGATGGC 7663
QY 179 ArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAlaIle 198
Db 7664 ATAGTAACAGCTGATATTATAGGAGTAGCCAAATGAAGATACGACGGGCTTGCTGCCATT 7723
QY 199 LeuAsnHis 201
Db 7724 CTCATAAT 7732

RESULT 6
AAH14096
ID AAH14096 standard; cDNA; 2157 BP.
XX
AC AAH14096;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:11265.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
Claim 8; SEQ ID NO 11265; 2537pp + Sequence Listing; English.
The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX
SQ Sequence 2157 BP; 544 A; 560 C; 591 G; 462 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.9e-59 Length: 2157
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 4 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x AAH14096 (1-2157)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 818 ATTGGATTTCATGGTGGCTGTATGACCCACTCACCAAAATTAATCCACTTTGGAGAAAG 877
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 878 GATTATGACATTTTGGCAGGCGGTGGACACACCTGACATAGAAATCTGGAAGAAT 937
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 938 ---GGGAAGGACCCAGCTCCTTTTAACTTGATCATGTTTAGGAATAACAACCTTGAAGC 994
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 995 AAAATCCATGACGTGAAGATTACATCACAGATGTTAACAGCTGGTGGTGACATTTGGT 1054
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 1055 TTCCATCTGCACAAATGCTATTCTCGATTCTCTGTTCCCAAAATTTGATTAAACAGAACCT 1114
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 1115 TCTTACGAACTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 1162
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 1163 TTCGGAGTCCAGCAGCAAGTGGCGCGGCGGCGGCAAGGCTTCTCTGCTGGCGGAAGATG 1222
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 1223 GCCGAGGTGCAG-----GTGAGCCGCGCGCGCGCGCGCGCGGCGGCGGCGGCTCTGG 1270
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 1271 CTGTGTTTCGCCACCGTCAAGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1330
QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
Db 1331 GGCGCGGTGCAGACCACTGCTCAACATCGCCACGAGGACTGCATCAAGGTGGCGGCC 1390
QY 198 IleLeuAsnHis 201
Db 1391 GTGCTCAACAAC 1402

RESULT 7
AAH14671
ID AAH14671 standard; cDNA; 3270 BP.
XX
XX
AC AAH14671;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12354.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN BP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 8; SEQ ID NO 12354; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to the polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX
SQ Sequence 3270 BP; 867 A; 821 C; 852 G; 730 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.49e-59
Score: 566.00
Percent Similarity: 71.08%
Best Local Similarity: 52.45%
Query Match: 52.55%
DB: 4
Length: 3270
Matches: 107
Conservative: 38
Mismatch: 47
Indels: 12
Gaps: 4

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

Claim 8; SEQ ID NO 11430; 2537pp + Sequence Listing; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the nucleotide sequence comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the polynucleotide comprises at least 15 nucleotides and the combination of oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

Sequence 3614 BP; 967 A; 898 C; 928 G; 821 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.03e-59	Length:	3614
Score:	566.00	Matches:	107
Percent Similarity:	71.08%	Conservative:	38
Best Local Similarity:	52.45%	Mismatches:	47
Query Match:	52.55%	Indels:	12
DB:	4	Gaps:	4

US-10-029-020-14_COPY_2400_2600 (1-201) x AAH14183 (1-3614)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
DB 2192 ATTGGATTTCATGGTGGCCCTGTATGACCCACTCACCAATTAATCCACTTTGGAGAAAGA 2251

QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40
DB 2252 GATTATGACATTTTGGCAGGACGGTGGACAACTGACATAGAAATCTGGAAAAGAATT 2311

QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
DB 2312 ---GGGAAGGACCCAGCTCCTTTTAACTTGTACATGTTAGGAATAACAACCTGCAAGC 2368

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
DB 2369 AAAATCCATGACGTGAAGATTACATCACAGATGTTAAACAGCTGGCTGGTGACATTGGT 2428

QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
DB 2429 TTCCATCTGCACAATGCTATTCTCTGGATTCCCTGTTCCCAAAATTGATTTAACAAGAACCT 2488

QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
DB 2489 TCTTACGAACCTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 2536

121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
2537 TTCGAGTCCAGCAGCAAGTGGCGGCGCAGGCCAAGGCCTTCTCTGTGCTGGGAAGATG 2596

141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
2597 GCCGAGGTGCAG-----GTGAGCCGCGCGCGCGCGCGCGCGCGCTCCTGG 2644

160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
2645 CTGTGGTTCGCCACGCTCAAGTCGTCATCGGCAAGGGCGTCTGTCGCGCTCAGCCAG 2704

178 GlyArgValThrAspIleIleSerValAlaAsnGluAspGlyArgValAlaAla 197
2705 GGCGCGGTGCAGACCAACGTGCTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGGCC 2764

198 IleLeuAsnHis 201
2765 GTGCTCAACAAC 2776

RESULT 9
ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX
AC ABQ82345;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15c encoding cDNA SEQ ID NO:39.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; infection; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; ss.

Homo sapiens.
Key Location/Qualifiers
CDS 258..8144
/*tag= a
/product= "NOV15c"

WO200262999-A2.
15-AUG-2002.
31-DEC-2001; 2001WO-US049976.
29-DEC-2000; 2000US-0258928P.
02-JAN-2001; 2001US-0259415P.
04-JAN-2001; 2001US-0259785P.
20-FEB-2001; 2001US-0269814P.
09-MAR-2001; 2001US-0279863P.
29-MAR-2001; 2001US-0279832P.
29-MAR-2001; 2001US-0279833P.
13-APR-2001; 2001US-0283889P.
18-APR-2001; 2001US-0284447P.
25-APR-2001; 2001US-0286683P.
29-MAY-2001; 2001US-0294080P.
16-AUG-2001; 2001US-0312915P.
17-AUG-2001; 2001US-0313325P.
17-SEP-2001; 2001US-0322699P.
26-NOV-2001; 2001US-0333350P.
(CURA-) CURAGEN CORP.

PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldov F, Burgess CE, Edinger S, Ellerman K;
XX Gunther E, Smithson G, Millet I, Macdougall JR;
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53588.

XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.

XX
PS Claim 8; Page 119-121; 444pp; English.

XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytotostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15c, which is
CC located on chromosome 4

SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.39e-58 Length: 8473
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 6 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x ABQ82345 (1-8473)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7050 ATTGGATTTCATGGTGGCCTGTATGACCCCACTACCAAAATTAATCCACTTTGGAGAAAGA 7109
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7110 GATTATGACATTTTGGCAGGACGGTGGCAACACCTGACATAGAAATCTGGAAGAATT 7169
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7170 ---GGGAAGACCCAGCTCCTTTTAACTTGTACATGTTTAGGAATAACAACCTGCAAGC 7226
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7227 AAAATCCATGACGTGAAAGATTACATACAGATGTTAACAGCTGGCTGGTACATTGGT 7286
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7287 TTCCATCTGCACAAATGCTATTCTCTGGATTCCCTGTTCCCAAAATTGATTTAACAGAACCT 7346
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120

Db 7347 TCTTACGAACTTGTG-----AAGAGTCAGCACTGGGATGATATACCGCCCATC 7394
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7395 TTCGGAGTCCAGCAGCAAGTGGCGGCAGGCCAAGGCCTTCCTGTGCTGGGGAAGATG 7454
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7455 GCCGAGGTGCAG-----GTGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7502
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7503 CTGTGTTTCGCCACGGTCAAGTCGCTGATCGGCAAGGGCGTCAATGTCGCGCTCAGCCAG 7562
QY 178 GlyArgValThrThrAspIleIleSerValAlaAlaAsnGluAspGlyArgValAlaAla 197
Db 7563 GGCGCGGTGCAGACCAACAGTGTCTCAACATCGCCAAACGAGGACTGCATCAAGGTGGCGCC 7622
QY 198 IleLeuAsnHis 201
Db 7623 GTGCTCAACAAC 7634
RESULT 10
ABQ82346
ID ABQ82346 standard; cDNA; 8487 BP.
XX ABQ82346;
AC
XX
DT 17-DEC-2002 (first entry)
XX Human NOV15d encoding cDNA SEQ ID NO:41.
DE
XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; goitre;
KW Pick's disease; vesicular transport disease; cystic fibrosis; allergy;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
Homo sapiens.
Key Location/Qualifiers
FT CDS 299..8140
FT /*tag= a
FT /product= "NOV15d"
XX
WO200262999-A2.
PD 15-AUG-2002.
XX
31-DEC-2001; 2001WO-US049976.
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.

PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shinkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53586.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 110-112; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytotatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4
XX
SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 1.44e-58 Length: 8675
Score: 566.00 Matches: 107
Percent Similarity: 71.08% Conservative: 38
Best Local Similarity: 52.45% Mismatches: 47
Query Match: 52.55% Indels: 12
DB: 6 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x ABQ82343 (1-8675)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7234 ATTGGATTTCATGGTGGCGCTGTATGACCCACTACCAAAATTAATCCACTTTGGAGAAAGA 7293
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7294 GATTATGACATTTTGGCAGGACGGTGGACACACCTGACATAGAAATCTGGAAAGAATT 7353
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProLeSer 60
Db 7354 ---GGGAAGGACCCAGCTCCTTTTAACTTGATACATGTTTAGGAATAACACCCCTGCAAGC 7410

QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7411 AAAATCCATGACGTGAAGATTACATCACAGATGTTAAACAGTGGTGGTGGTGGTGGT 7470
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7471 TTCCATCTGCACAATGCTATTCTCTGGATTCCCTGTTCCCAAATTTGATTAAACAGAACCT 7530
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTyrAspAsnSerLysSerIle 120
Db 7531 TCTTACGAACCTGTG-----AAGAGTCAGCAGTGGGATGATATACCGCCCATC 7578
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7579 TTCGGAGTCCAGCAGCAAGTGGCGGCGGCAAGGCCCTTCTGTGCTGGGGAAGATG 7638
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7639 GCCGAGGTGCAG-----GTGAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7686
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAsp 177
Db 7687 CTGTGGTTCGCCACGTCGAAGTCGCTGATCGCGCAAGGCGCGTCATGTGCGCGTCAGCCAG 7746
QY 178 GlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAla 197
Db 7747 GGCGCGGTGCAGACCAACGTCGTCACATCGCCCAACGAGGACTGCATCAAGTGGCGCGCC 7806
QY 198 IleLeuAsnHis 201
Db 7807 GTGCTCAACAAC 7818
RESULT 13
ABK34316
ID ABK34316 standard; cDNA; 791 BP.
XX
AC ABK34316;
XX
DT 08-MAY-2002 (first entry)
XX
DE Human cDNA for novel secreted protein, SEQ ID 85.
XX
KW Human; ss; gene; secreted protein; immune deficiency; viral infection;
KW bacterial infection; fungal infection; autoimmune disorder; burn;
KW rheumatoid arthritis; multiple sclerosis; autoimmune thyroiditis;
KW diabetes; allergy; asthma; neurodegenerative disease; Crohn's disease;
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; tumour;
KW coagulation disorder; haemophilia; inflammatory disorder; ulcer;
KW tissue regeneration; wound healing; haematopoiesis; myeloid deficiency;
KW lymphoid cell deficiency.
XX
OS Homo sapiens.
XX
PN WO200177290-A2.
XX
PD 18-OCT-2001.
XX
PF 29-MAR-2001; 2001WO-US010295.
XX
PR 06-APR-2000; 2000US-0194941P.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
PI Gulukota K, Graham JR;
XX
XX WPI; 2002-179323/23.
DR
XX
PT Six hundred and twenty five polynucleotides derived from a variety of
PT human tissue sources which encode secreted proteins, useful for treating
PT immune deficiencies and disorders such as autoimmune disorders.
XX

PS Claim 1; Page 91-92; 339pp; English.
CC The invention relates to 625 polynucleotides which have been derived from
CC a variety of human tissue sources and which encode novel secreted
CC proteins, their complements and sequences that hybridize to them. Also
CC included are a vector comprising the polynucleotide, a host cell
CC transformed with the vector, the proteins encoded by the polynucleotides,
CC antibodies that bind to the proteins and identification of modulators of
CC the proteins or the expression of the polynucleotide. The polynucleotides
CC can be used as probes for the identification and isolation of full length
CC cDNA and genomic DNA. The polynucleotides and proteins can also be used
CC as nutritional supplements. The protein is useful in the treatment of
CC various immune deficiencies and disorders such as viral infections,
CC bacterial infections, fungal infections, autoimmune disorders (e.g.
CC rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis and
CC diabetes) and allergic reactions and conditions (e.g. asthma). They are
CC also useful for treating neurodegenerative diseases (e.g. Alzheimer's
CC disease, Parkinson's disease), liver fibrosis, coagulation disorders
CC (e.g. haemophilia), inflammatory disorders (e.g. Crohn's disease) and
CC tumours. They are also useful for tissue regeneration, for wound healing
CC and in the treatment of burns, incisions and ulcers. The proteins are
CC also useful for regulating haematopoiesis, for treating myeloid or
CC lymphoid cell deficiencies. The present sequence is one of the 625 cDNA
CC sequences encoding a secreted protein
XX

SQ Sequence 791 BP; 200 A; 195 C; 209 G; 183 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 7.66e-59 Length: 791
Score: 556.00 Matches: 106
Percent Similarity: 71.14% Conservative: 37
Best Local Similarity: 52.74% Mismatches: 52
Query Match: 51.62% Indels: 7
DB: 6 Gaps: 3

US-10-029-020-14_COPY_2400_2600 (1-201) x ABK34316 (1-791)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 102 ATTGGATTTCATGGTGGCCTGTATGACCCACTCACCAATTATCCACTTTGGAGAAAGA 161
QY 21 AspTyrAspValLeuAlaGlyArgTrpThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 162 GATTATGACATTTTGGCAGGACGGTGGACACACCTGACATAGAAATCTGGAAAGAAT 221
QY 41 SerSerSerAsnValMetProPheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 222 ---GGGAAGGACCCAGCTCTTTAACTTGATGTTAGGAATAACAACCTGCAAGC 278
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 279 AAATCCATGACGTGAAGATTACATCACAGATGTTAACAGCTGGCTGGTGACATTGGT 338
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 339 TTCCATCTGCACAATGCTATTCTTGATTTCCTGTCCCAATTGTGATTTAACAGAACCT 398
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 399 TCTTACGAACCTTGTG-----AAGAGTCAGCAGTGGGATGATATACCGTCCATC 446
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 447 TTCGGAGTCCAGCAGCAAGTGGCGCGCAGGCCAAGGTCTTCCTGTGCTGGGGAAGATG 506
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLysPhe 160
Db 507 GCCGAGGTGCAGGTGAGCCGGCGCGCGCGCGCGAGTCCTGCTGTGG---TTC 562
QY 161 AlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLysAspGlyArgVal 180
Db 563 GCCNCGGTCAAGTCGCTGATCGGCAAGGGGCTCATGCTGGCNGTCAGCCAGGCCGCGTG 622

QY 181 ThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAlaAlaIleLeuAsn 200
Db 623 CAGACCAACGTGTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGCGTGTCAAC 682
QY 201 His 201
Db 683 AAC 685
RESULT 14
ACC72051
ID ACC72051 standard; DNA; 9058 BP.
XX
AC ACC72051;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205A gene #SEQ ID 79.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
PN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58317.
XX

New isolated human differentially-regulated breast cancer polynucleotide and polypeptide, useful for diagnosing, staging, prognosticating, preventing and/or treating diseases and conditions relating to breast cancer.

Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.

The invention relates to isolated polynucleotides which are differentially-regulated in breast cancer. The methods and compositions of the present invention are useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing and/or treating diseases and conditions relating to breast cancer, and may be used in gene therapy or antisense therapy. They can also be used in research, drug discovery, clinical medicine and forensic medicine. Sequences given in records ACC72012-ACC72074 represent polynucleotides of the invention that are differentially-regulated in breast cancer. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 5.1e-54 Length: 9058
Score: 529.50 Matches: 103
Percent Similarity: 68.78% Conservative: 38
Best Local Similarity: 50.24% Mismatches: 49
Query Match: 49.16% Indels: 15
DB: 7 Gaps: 4

US-10-029-020-14_COPY_2400_2600 (1-201) x ACC72051 (1-9058)

QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20

Db 6599 ATTGGCTTCCATGGGGACTCTATGACCCCTGACCAAGCTGGTCCACTTCACTCAGCGT 6658
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 6659 GATTATGATGTGTGTCAGGACGATGGACCTCCCGAGACTATACCATGTGGAAAAACGTG 6718
QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 6719 GGCAAGGAGCCGGCC---CCCTTTAACCTGTATATGTTCAAGAGCAACAATCCTCTCAGC 6775
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 6776 AGTGAGCTAGATTGAAGAACTACGTGACAGATGTGAAAAGCTGGCTTGTGATGTTTGA 6835
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 6836 TTTCAGCTTAGCAACATCATTCCTGGCTTCCCGAGAGCCAAAATGTATTTCGTGCTCCT 6895
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 6896 CCCTATGAATTGTCAGAGATCAAGCAAGT-----GAGAATGGACAGCTCATT 6943
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 6944 ACAGGTGTCCAAACAGACACAGAGAGACATAACCGGCCTTCATGGCTCTGGAA----- 6997
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 6998 -----GGACAGGTCAATTACTAAAAGCTCCACGCCAGCATCCCGAGAGAAAGCA 7045
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys 176
Db 7046 GGTCACCTGGTTTGCCACCACCACCGCCCATCATTTGGCAAAGGCATCATGTTGCCATCAA 7105
QY 177 AspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAla 196
Db 7106 GAAGGGCGGTGACCACCGGCGTGTCTCCAGCATCGCCAGCGAAGATAGCCCGAAGGTGGCA 7165
QY 197 AlaIleLeuAsnHis 201
Db 7166 TCTGTGCTGAACAAC 7180
RESULT 15
ID ACC72052 standard; DNA; 9695 BP.
XX AC ACC72052;
XX 08-JUL-2003 (first entry)
DT BCU0205B gene #SEQ ID 81.
XX Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX Homo sapiens.
OS WO2003029421-A2.
XX 10-APR-2003.
XX 02-OCT-2002; 2002WO-US031287.
XX 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
PA Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX WPI; 2003-381623/36.

DR P-PSDB; ABR58318.
XX New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
PS The invention relates to isolated polynucleotides which are
XX differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.63e-54 Length: 9695
Score: 529.50 Matches: 103
Percent Similarity: 68.78% Conservative: 38
Best Local Similarity: 50.24% Mismatches: 49
Query Match: 49.16% Indels: 15
DB: 7 Gaps: 4
US-10-029-020-14_COPY_2400_2600 (1-201) x ACC72052 (1-9695)
QY 1 IleGlyTyrHisGlyGlyLeuTyrAspProLeuThrLysLeuValHisMetGlyArgArg 20
Db 7236 ATTGGCTTCCATGGGGACTCTATGACCCCTGACCAAGCTGGTCCACTTCACTCAGCGT 7295
QY 21 AspTyrAspValLeuAlaGlyArgTyrThrSerProAspHisGluLeuTrpLysHisLeu 40
Db 7296 GATTATGATGTGTGTCAGGACGATGGACCTCCCGAGACTATACCATGTGGAAAAACGTG 7355
QY 41 SerSerSerAsnValMetPropheAsnLeuTyrMetPheLysAsnAsnProIleSer 60
Db 7356 GGCAAGGAGCCGGCC---CCCTTTAACCTGTATATGTTCAAGAGCAACAATCCTCTCAGC 7412
QY 61 AsnSerGlnAspIleLysCysPheMetThrAspValAsnSerTrpLeuLeuThrPheGly 80
Db 7413 AGTGAGCTAGATTGAAGAACTACGTGACAGATGTGAAAAGCTGGCTTGTGATGTTTGA 7472
QY 81 PheGlnLeuHisAsnValIleProGlyTyrProLysProAspMetAspAlaMetGluPro 100
Db 7473 TTTCAGCTTAGCAACATCATTCCTGGCTTCCCGAGAGCCAAAATGTATTTCGTGCTCCT 7532
QY 101 SerTyrGluLeuIleHisThrGlnMetLysThrGlnGluTrpAspAsnSerLysSerIle 120
Db 7533 CCCTATGAATTGTCAGAGATCAAGCAAGT-----GAGAATGGACAGCTCATT 7580
QY 121 LeuGlyValGlnCysGluValGlnLysGlnLeuLysAlaPheValThrLeuGluArgPhe 140
Db 7581 ACAGGTGTCCAAACAGACACAGAGAGACATAACCGGCCTTCATGGCTCTGGAA----- 7634
QY 141 AspGlnLeuTyrGlySerThrIleThrSerCysGlnGlnAlaProLysThrLysLys--- 159
Db 7635 -----GGACAGGTCAATTACTAAAAGCTCCACGCCAGCATCCCGAGAGAAAGCA 7682
QY 160 -----PheAlaSerSerGlySerValPheGlyLysGlyValLysPheAlaLeuLys 176
Db 7683 GGTCACTGGTTTGCCACCACCGCCCATCATTTGCAAAAGGCATCATGTTTGCATCAAA 7742
QY 177 AspGlyArgValThrThrAspIleIleSerValAlaAsnGluAspGlyArgArgValAla 196
Db 7743 GAAGGGCGGTGACCACGGCGGTGTCTCCAGCATCGCCAGCGAAGATAGCCCGAAGGTGGCA 7802

Qy 197 AlarLeLeuAsnHis 201
Db :::::|||||:::
7803 TCTGTGCTGAACAAC 7817

Search completed: August 14, 2004, 02:34:36
Job time : 386.95 secs

This Page Blank (uspio)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 1084.27 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
Perfect score: 2845
Sequence: 1 YYIGADGSLRLLLANGMEVA.....YNRAGSWSVRYDGLGRRV 541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn_1_1_2156@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description

1	2845	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
2	2845	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
3	2821.5	99.2	8438	13	US-10-042-865-1	Sequence 1, Appli
4	2751	96.7	8355	13	US-10-383-201-55	Sequence 55, Appl
5	2233	78.5	3614	13	US-10-342-887-1743	Sequence 1743, Ap
6	2233	78.5	3614	13	US-10-172-118-1743	Sequence 1743, Ap
7	2233	78.5	8473	17	US-10-038-854-39	Sequence 39, Appl
8	2233	78.5	8487	17	US-10-038-854-41	Sequence 41, Appl
9	2233	78.5	8645	17	US-10-038-854-37	Sequence 37, Appl
10	2233	78.5	8675	17	US-10-038-854-35	Sequence 35, Appl
11	2141.5	75.3	8797	9	US-09-808-602-74	Sequence 74, Appl
12	2141.5	75.3	8797	9	US-09-808-602-77	Sequence 77, Appl
13	2141.5	75.3	8797	10	US-09-800-198-62	Sequence 62, Appl
14	2141.5	75.3	8797	10	US-09-800-198-65	Sequence 65, Appl
15	2138.5	75.2	8409	9	US-09-808-602-79	Sequence 79, Appl
16	2138.5	75.2	8409	10	US-09-800-198-67	Sequence 76, Appl
17	2136.5	75.1	6560	9	US-09-808-602-76	Sequence 64, Appl
18	2136.5	75.1	6560	10	US-09-800-198-64	Sequence 79, Appl
19	2136.5	75.1	9058	16	US-10-144-194A-79	Sequence 81, Appl
20	2136.5	75.1	9695	16	US-10-144-194A-81	Sequence 12, Appl
21	2136.5	75.1	9729	9	US-09-808-602-12	Sequence 12, Appl
22	2136.5	75.1	9729	10	US-09-800-198-12	Sequence 7, Appli
23	2136.5	75.1	9826	9	US-09-800-198-7	Sequence 78, Appl
24	2136.5	75.1	8689	9	US-09-808-602-78	Sequence 66, Appl
25	2135.5	75.1	8689	10	US-09-800-198-66	Sequence 143, App
26	2135.5	75.1	8689	13	US-10-072-012-143	Sequence 927, App
27	2106.5	74.0	8575	13	US-10-295-027-927	Sequence 21273, A
28	2048.5	72.0	12880	16	US-09-864-761-21273	Sequence 9717, Ap
c 29	982.5	34.5	806	9	US-10-198-846-9717	Sequence 4526, Ap
c 30	982.5	34.5	1207	15	US-09-864-761-4526	Sequence 7548, Ap
c 31	982.5	34.5	1973	9	US-10-198-846-7548	Sequence 13976, A
32	735.5	25.9	648	15	US-10-198-846-13976	Sequence 6607, Ap
33	665	23.4	3217	15	US-09-864-408A-6607	Sequence 104, App
34	520	18.3	330	11	US-10-040-739-104	Sequence 15932, A
35	387	13.6	286	14	US-10-029-386-15932	Sequence 2232, Ap
36	369	13.0	307	15	US-10-369-493-34930	Sequence 34930, A
37	369	13.0	307	15	US-10-029-386-2232	Sequence 3302, Ap
38	248.5	8.7	3957	16	US-10-282-122A-11416	Sequence 11416, A
39	247	8.7	586	15	US-10-029-386-3302	Sequence 17002, A
40	240.5	8.5	4545	13	US-10-029-386-17002	Sequence 32730, A
41	240	8.4	245	15	US-10-282-122A-32730	Sequence 32058, A
42	219	7.7	3759	13	US-10-282-122A-32058	Sequence 3837, Ap
43	212	7.5	4155	13	US-10-282-122A-32058	Sequence 33231, A
c 44	205	7.2	5394	16	US-10-398-221-3837	
45	204.5	7.2	4716	13	US-10-282-122A-33231	

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)

```

Alignment Scores:		
Pred. No.:	0	8354
Score:	2845.00	541
Percent Similarity:	100.00%	
Best Local Similarity:	100.00%	
Query Match:	100.00%	
DB:	13	
	Length:	
	Matches:	
	Conservative:	0
	Mismatches:	0
	Indels:	0
	Gaps:	0

US 10-038-030-14 COPY 1760 2300 (1-541) X US-10-383-201-43 (1-8354)

QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla	20
Db	5312	TACTACATCGGGCCGATGGCTCCTTTCGGCTGCTGTGGCCAACGGCATGGAGGTGGCG	5371
QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn	40
Db	5372	CTGCAGACTGAGCCCACTTGCTGGCTGGCACCGTCAACCCCAACCGTGGCAAGAGGAAT	5431
QY	41	ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln	60
Db	5432	GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCCAAAGAGCAG	5491
QY	61	AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu	80
Db	5492	GCTCGGGGCCAGGTCACTGTCTTTGGCGCCGGCTGCGGGTGCACAACCGAAATCTCCTA	5551
QY	81	SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe	100
Db	5552	TCTCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATGATGACCCACCGCAAGTTC	5611
QY	101	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg	120
Db	5612	ACCTTCGGATTCTGTACGACCAGGGGGGGCGGCCAGCCTCTGGTCACCCAGCAGCAGG	5671
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly	140
Db	5672	CTGAATGGTGTCAACGTGACATACTCCCTGGGGGTATCATTTGCTGGCATCCAGAGGGGC	5731
QY	141	IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla	160
Db	5732	ATCATGTCTGAAAGAAATGGAATACGACCAGGGGGCCGCATCACATCCAGGATCTTCGCT	5791
QY	161	AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer	180
Db	5792	GATGGGAAGACATGGAGCTACACATACTTAGAGAAGTCCATGGTGTCTACTACACAGC	5851
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro	200
Db	5852	CAGAGGCAGTATATCTTTGAGTTCGACAAGAAATGACCGCCTCTCTTCTGTGACGATGCC	5911
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr	220
Db	5912	AACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTAT	5971
QY	221	GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu	240

5972	Db		CAGCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGGACCTCCTT	6031
241	QY		HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys	260
6032	Db		CACACCTTCTACCTGGGCACTGGCCGACGGGTGATATACAAGTATGGCAAACCTCAAAG	6091
261	QY		LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly	280
6092	Db		CTGGCAGAGACGCTCTATGACACCAACCAAGGTCAGTTTCACTATGACGAGACGCGAGGC	6151
281	QY		MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln	300
6152	Db		ATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACTGACCATCCGCTACCGTCAG	6211
301	QY		IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGlyMetValAsnAla	320
6212	Db		ATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCC	6271
321	QY		ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn	340
6272	Db		CGTTTTGACTACAACTATGACCAACAGCTTCCGGTGATGATGTCAGGCAAGACAGAGCAG	6331
341	QY		GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln	360
6332	Db		GAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTCAGGCAAGACAGAGCAG	6391
361	QY		PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet	380
6392	Db		TTTGGGAAGTTGGTGTCAATTACTATGACATTAAACCAGATCATCACCACAGCTGTCAATG	6451
381	QY		ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
6452	Db		ACCCACACCAAGCATTTTGATGCATATGGCAGGATGAAGGAAGTCAGTATGAGATCTTC	6511
401	QY		ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420
6512	Db		CGCTCGCTCATGTACTGGATGACCGCTCCAGTATGATAACATGGGGCGAGTAGTGAAGAAG	6571
421	QY		GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
6572	Db		GAGCTGAAGGTAGACCCCTACGCCAATACCACTCGCTACTCTCTATGATGATGCTGAC	6631
441	QY		GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu	460
6632	Db		GGCCAGCTGCAGACAGTCTCCATCAATGACAAAGCCACTCTGGCGCTACAGCTACGACCTC	6691
461	QY		AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
6692	Db		AATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCAGGCTCACACCACTACGGTAT	6751
481	QY		AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
6752	Db		GACATCCGCGACCGCATCACTCGCTGGGTGACGTGCAATACAAGATGGATGAGGATGGC	6811
501	QY		PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
6812	Db		TTCTGTAGGCAGCGGGCGGTGATATCTTTTGAGTACAACTCAGCTGGCCTGCTCATCAAG	6871
521	QY		AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
6872	Db		GCCTACAACCGGGCTGGCAGTGGAGTGTCAAGTACCGCTACGATGGCCTGGGGCGGCGC	6931
541	QY		Val 541	
6932	Db		GTG 6934	

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US2004003397A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.

; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-029-020-13

; Alignment Scores:
; Pred. No.: 0
; Score: 2845.00
; Percent Similarity: 100.00%
; Best Local Similarity: 100.00%
; Query Match: 100.00%
; DB: 13

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
1	2845.00	8354	541	0	0	0	0
5312	100.00%						
21	100.00%						
5372	100.00%						
41	100.00%						
5432	100.00%						
61	100.00%						
5492	100.00%						
81	100.00%						
5552	100.00%						
101	100.00%						
5612	100.00%						
121	100.00%						
5672	100.00%						
141	100.00%						
5732	100.00%						
161	100.00%						

Db	5792	GATGGGAAGACATGGAGCTACACATACTTAGAGAAATCCATGGTGTGCTACTACACAGC	5851
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro	200
Db	5852	CAGAGGCAGTATATCTTTGAGTTCGACAAGAAATGACCGCTCTCTTCTGTGACGATGCC	5911
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr	220
Db	5912	AACGTGGCGGGCAGACACTAGAGACCATCCGCTCAGTGGCTACTACAGAAACATCTAT	5971
QY	221	GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu	240
Db	5972	CAGCCCCCTGAGGCAATGCCTCAGTCATACAGGACTTCACCTGAGGATGGGACCTCCTT	6031
QY	241	HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys	260
Db	6032	CACACCTTCTACCTGGGCACCTGGCCGCGAGGTGATATACAAGTATGGCAAACTGTCAAAG	6091
QY	261	LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly	280
Db	6092	CTGGCAGAGACGCTCTATGACACCACCAAGGTGAGTTTACCTATGACGAGACGGCAGGC	6151
QY	281	MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln	300
Db	6152	ATGCTGAAGACCATCAACCTACAGAAATGAGGGCTTCACTGCACCATCCGCTACCGTCAG	6211
QY	301	IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluMetValAsnAla	320
Db	6212	ATTGGGCCCCGATGATGACCGACAGATCTTCCGCTTCACTGAGGAAGCATGGTCAACGCC	6271
QY	321	ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn	340
Db	6272	CGTTTGACTACAACTATGACAACAGCTTCCGGGTGACCAGCATGCAGGCTGTGATCAAC	6331
QY	341	GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln	360
Db	6332	GAGACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGTCAGGCAAGACAGAGCAG	6391
QY	361	PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet	380
Db	6392	TTTGGGAAGTTTGGTGTCAITTAATGATGACATTAACCAAGATCATCACACAGCTGTCAIT	6451
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
Db	6452	ACCCACACCAAGCATTTTGTATGCATATGGCAGGATGAAGAAAGTGCAGTATGAGATCTTC	6511
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420
Db	6512	CGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAAACAATGGGCGAGTAGTGAAGAAG	6571
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
Db	6572	GAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCCTATGAGTATGATGCTGAC	6631
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu	460
Db	6632	GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGGTACAGCTACGACCTC	6691
QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
Db	6692	AATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTAT	6751
QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
Db	6752	GACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGATGGC	6811
QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
Db	6812	TTCTTGAGGACGGGGCGGTGATATCTTTGAGTACAACACTCAGCTGGCCCTGCTCATCAAG	6871
QY	521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540

Db	6872	GCCTACAACGGGCTGCACAGCTGGAGTGTACGGTACCGCTACGATGGCCTGGGGGGCGGC	6931
QY	541	Val 541	
Db	6932	GTG 6934	
RESULT 3			
US-10-042-865-1			
; Sequence 1, Application US/10042865			
; Publication No. US20040029216A1			
; GENERAL INFORMATION:			
; APPLICANT: Padigaru, Muralidhara			
; APPLICANT: Li, Li			
; APPLICANT: Zerhusen, Bryan D			
; APPLICANT: Casman, Stacie J			
; APPLICANT: Shenoy, Suresh G			
; APPLICANT: Spytek, Kimberly			
; APPLICANT: Zhong, Mei			
; APPLICANT: Gangolli, Esha A			
; APPLICANT: Burgess, Catherine E			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Vernet, Corine A.M			
; APPLICANT: Taylor, Sarah			
; APPLICANT: Tchernev, Velizar T			
; APPLICANT: Miller, Charles E			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Boldog, Ference L			
; APPLICANT: Grosse, William M			
; APPLICANT: Alsobrook II, John P			
; APPLICANT: Gerlach, Valerie L			
; APPLICANT: Edinger, Shlomit R			
; APPLICANT: Rothenberg, Mark E			
; APPLICANT: Ellerman, Karen			
; APPLICANT: MacDougall, John			
; APPLICANT: Malyankar, Uriel M			
; APPLICANT: Millet, Isabelle			
; APPLICANT: Peyman, John			
; APPLICANT: Smithson, Glennda			
; APPLICANT: Gunther, Erik			
; APPLICANT: Stone, David			
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of			
; FILE REFERENCE: 21402-537			
; CURRENT APPLICATION NUMBER: US/10/042,865			
; CURRENT FILING DATE: 2002-05-17			
; PRIOR APPLICATION NUMBER: 60/260,417			
; PRIOR FILING DATE: 2001-01-09			
; PRIOR APPLICATION NUMBER: 60/260,831			
; PRIOR FILING DATE: 2001-01-10			
; PRIOR APPLICATION NUMBER: 60/272,338			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: 60/274,876			
; PRIOR FILING DATE: 2001-03-09			
; PRIOR APPLICATION NUMBER: 60/284,704			
; PRIOR FILING DATE: 2001-04-18			
; NUMBER OF SEQ ID NOS: 264			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1			
; LENGTH: 8438			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-042-865-1			
Alignment Scores:			
Pred. No.:	0	Length:	8438
Score:	2821.50	Matches:	541
Percent Similarity:	98.72%	Conservative:	0
Best Local Similarity:	98.72%	Mismatches:	0
Query Match:	99.17%	Indels:	7
DB:	13	Gaps:	2
US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-042-865-1 (1-8438)			

QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla	20
Db	5318	TACTACATCGGGGCCGATGGCTCCTTGGCGTGTCTGGCCACCGCATGGAGTGGCG	5377
QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn	40
Db	5378	CTGCAGACTGAGCCCCACTTGTGTGGTGGCACCGTCAACCCCCACCGTGGGCAAGAGGAAT	5437
QY	41	ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln	60
Db	5438	GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAG	5497
QY	61	AlaArgGlyGlnValThrValPheGlyArgArgLeuArg-----ValHisAsnArg	77
Db	5498	GCTCGGGGCCAGGTCACTGTCTTTGGCGCGCGGCTGCGGTTGCTCCAGGTTTCAACCCGA	5557
QY	78	AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHis	97
Db	5558	AATCTCTATCTCTGGACTTTGTATCGCGTAACACGCACAGAGAAGATCTATGATGACCAC	5617
QY	98	ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerPro	117
Db	5618	CGCAAGTTCAACCTTGGATTCTGTACGACCAGCGGGGGCGCCAGCCTCTGTGTACCC	5677
QY	118	SerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIle	137
Db	5678	AGCAGCAGGCTGAATGGTGTCAACGTGACATACCTCCCTGGGGGTACATTGCTGGCATC	5737
QY	138	GlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArg	157
Db	5738	CAGAGGGGCATCATGTCTGAAAGAAATGGAATACGACCAGCGGGCGGCATCACATCCAGG	5797
QY	158	IlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLys-----Ser	173
Db	5798	ATCTTCGCTGATGGGAAGACATGGAGCTACACATACCTTAGAGAAGGAGGTGTCCAGTCC	5857
QY	174	MetValLeuLeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArg	193
Db	5858	ATGGTGTCTGTACTACACAGCCAGAGGAGTATATCTTTGAGTTTCGACAAGAATGACCGC	5917
QY	194	LeuSerSerValThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerVal	213
Db	5918	CTCTCTTCTGTGACGATGCCCAACGTGGCGGGCAGACACTAGAGACCATCCGCTCAGTG	5977
QY	214	GlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPhe	233
Db	5978	GGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTTCATACAGGACTTC	6037
QY	234	ThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyr	253
Db	6038	ACTGAGGATGGGCACCTCCTTACACCTTCTACCTGGGCACCTGGCGGCGGTGATATAC	6097
QY	254	LysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPhe	273
Db	6098	AAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCAGTTTC	6157
QY	274	ThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThr	293
Db	6158	ACCTATGACGAGCGGCGAGCATGTGAAGACCATCAACCTACAGAAATGAGGGCTTCACC	6217
QY	294	CysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThr	313
Db	6218	TGCACCATCCGCTACCGTCAGATTGGGCCCCCTGATTGACCGACAGATCTTCCGCTTCACT	6277
QY	314	GluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThr	333
Db	6278	GAGGAAGGCATGGTCAACGCCCGTTTTTGACTACATATGACAACAGCTTCCGGGTGACC	6337
QY	334	SerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAsp	353
Db	6338	AGCATGCAGGCTGTGATCAACAGACCCCCACTGCCCATTTGATCTCTATCGCTATGATGAT	6397
QY	354	ValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGln	373

```
Db 6398 GTGTGAGGCAAGACAGACAGTGTGGAAAGTTTGGTGTCAATTTACTATGACATTAAACAG 6457
QY 374 IleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLys 393
Db 6458 ATCATCACCACAGCTGTATGACCCACACCAAGCATTTTGATGCATATGGCAGGATGAAG 6517
QY 394 GluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsn 413
Db 6518 GAAAGTGAGTATGATGATCTTCGGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAAC 6577
QY 414 MetGlyArgValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyr 433
Db 6578 ATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACCTCGCTAC 6637
QY 434 SerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeu 453
Db 6638 TCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTC 6697
QY 454 TrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAla 473
Db 6698 TGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCA 6757
QY 474 ArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGln 493
Db 6758 CGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTCAA 6817
QY 494 TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn 513
Db 6818 TACAAGATGGATGAGGATGGCTTCCTGAGGCAGCGGGCGGTGATATCTTTGAGTACAAC 6877
QY 514 SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArg 533
Db 6878 TCAGCTGGCCTGCTCATCAAGGCCTACAACCGGCTGGCAGCTGGAGTGTGAGTACCGC 6937
QY 534 TyrAspGlyLeuGlyArgArgVal 541
Db 6938 TACGATGGCTGGGGCGGCGGTG 6961

RESULT 4
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 0
Score: 2751.00
Percent Similarity: 97.24%
Best Local Similarity: 97.06%
Query Match: 96.70%
DB: 13

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-383-201-55 (1-8355)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db 5296 TACTACATCGGGCCGATGGCTCCTTGGCGGTGCTGCTGGCCAAACGGCATGGAGGTGGCG 5355
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5356 CTGCAGACTGAGCCCACTTGTGGCTGGCACCGTCAACCCCAACCGTGGCAAGAGGAAT 5415
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5416 GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCACGCAAGAGCAG 5475
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArg-----ValHisAsnArg 77
Db 5476 GCTCGGGCCAGGTCACTGTCTTTGGGGCCGGCTGCGGGTGTCCAGGTTCAACAACCGA 5535
QY 78 AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHis 97
Db 5536 AATCTCCTATCTCTGGACTTTGATCGCGTAACACGCACAGAGAGATCTATGATGACCAC 5595
QY 98 ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerPro 117
Db 5596 CGCAAGTTCACCCCTTCGGATTCTGTACGACCGAGCGGGCGGCCAGCCTCTGGTCACCC 5655
QY 118 SerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIle 137
Db 5656 AGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTGGGGTTACATTGCTGGCATC 5715
QY 138 GlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArg 157
Db 5716 CAGAGGGGCATCATGTCTGAAGAATGGAATACGACCAGCGGGCGGCATCACATCCAGG 5775
QY 158 IlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeu 177
Db 5776 ATCTTCGCTGATGGGAAGACATGGAGCTACACATACTATTAGAGAAGGCGAGGTGTC----- 5829
QY 178 LeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerVal 197
Db 5830 -----GAGTTTCGACAAGAATGACCGCCTCTCTTCTGTG 5862
QY 198 ThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArg 217
Db 5863 ACGATGCCCAACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGA 5922
QY 218 AsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGly 237
Db 5923 AACATCTATCAGCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACCTGAGGATGGG 5982
QY 238 HisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLys 257
Db 5983 CACCTCCTTACACCTTCTACCTGGGCACCTGGCCGACGGTGTATATACAAGTATGGCAA 6042
QY 258 LeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGlu 277
Db 6043 CTGTCAAAGCTGGCAGAGACGCTCTATGACACCACCAAGGTCAAGTTTACCTATGACGAG 6102
```


Db 992 CAAACAGCTTTCTTGGGTACAAAGTCGGAGGGTCTTATTCAAATACAGAAGCGACTAGG 1051

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
||||:|||||
Db 1052 CTCTCAGAAATTTTATATGATAGACACAAGAGTCAGTTTACCTATGATGAACAGCAGGA 1111

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
:::|||||
Db 1112 GTCCTAAAGACAGTAACCTCCAGAGTGATGGTTTATTTCACCATTAGATACAGGCAA 1171

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
|||||
Db 1172 ATTGGTCCCTGATTGACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGGTAAATGCA 1231

QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
|||||
Db 1232 AGATTTGACTATAGCTATGACAACAGCTTTCCAGTGACCAGCATGCAGGGTGTGATCAAT 1291

QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
|||||
Db 1292 GAAACGCCACTGCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG 1351

QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
|||||
Db 1352 TTTGGAAAGTTGGAGTTATATATATGATATTAACCAAGATCAATTTCTACAGCTGTAATG 1411

QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
|||||
Db 1412 ACCTATACGAAGCACTTTGATGTCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC 1471

QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
|||||
Db 1472 AGGTCGCTCATGTACTGGATTACAAATTCAGTATGATAACATGGGTCGGGTAACCAAGAGA 1531

QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
|||||
Db 1532 GAGATTAAATAGGGCCCTTTTGCCAAACACCACCAATATGCTTATGAATATGATGTTGAT 1591

QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
|||||
Db 1592 GGACAGCTCCAAACAGTTTACCTCAATGAAAGATAATGTGCGCGGTACAACACTACGATCTG 1651

QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
|||||
Db 1652 AATGAAACCTCCATTTACTGAACCAAGTAAACAGTGC CGCTGTGACACCCCTTCGCTAT 1711

QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
|||||
Db 1712 GACCTCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT 1771

QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
|||||
Db 1772 TTCCTACGTCAAAGGGGCACGAAATCTTTGAATATAGCTCAAGGGGCTTCTAACTCGA 1831

QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
|||||
Db 1832 GTTTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCGTTATGACGGCCTTGGGAAGGCGT 1891

QY 541 Val 541
|||

Db 1892 GTT 1894

RESULT 6

US-10-172-118-1743

; Sequence 1743, Application US/10172118

; Publication No. US20030224374A1

; GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue

; APPLICANT: He, Yudong

; APPLICANT: Linsley, Peter

; APPLICANT: Mao, Mao

; APPLICANT: Roberts, Chris

; APPLICANT: Van 't Veer, Laura

; APPLICANT: Van de Vijver, Marc

; APPLICANT: Bernards, Rene

; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-175-999

; CURRENT APPLICATION NUMBER: US/10/172,118

; CURRENT FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 60/380,770

; PRIOR FILING DATE: 2002-05-14

; NUMBER OF SEQ ID NOS: 2699

; SEQ ID NO 1743

; LENGTH: 3614

; TYPE: DNA

; ORGANISM: Homo sapiens

; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NM_018104

; DATABASE ENTRY DATE: 2001-06-18

US-10-172-118-1743

Alignment Scores:

Pred. No.: 2.88e-257 Length: 3614

Score: 2233.00 Matches: 404

Percent Similarity: 88.91% Conservative: 77

Best Local Similarity: 74.68% Mismatches: 60

Query Match: 78.49% Indels: 0

DB: 13 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-172-118-1743 (1-3614)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
|||
Db 272 TACCAGATTGGTTATGACGGCTCCCTCAGAATTATCTACGCCAGTGGCCTGGACTCACAC 331

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
|||||
Db 332 TACCAAAACAGAGCCGCACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAAAC 391

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
:::|||||
Db 392 ATGACTTTGCCCTGGCGAGAACGGTCAAAAACCTTGGTGGAAATGGAGATTCCGAAAGAGCAA 451

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
|||||
Db 452 GCCCAAGGGAAGTCAATGTCTTTGGCCGAAGCTCAGGGTTAATGGCAGAAACCTCCTT 511

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
|||||
Db 512 TCAGTTGACTTTGATCGAAACAACAAAGACAGAAAAGATCTATGACGACCACCGTAAATT 571

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
|||||
Db 572 CTA CTGATGAGGATCGCCTACGACACGCTCTGGGCACCCCGACTCTCTGGCTGCCAAGCAGCAAG 631

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
|||||
Db 632 CTGATGCGCGTCAATGTTCACCTATTTCATCCACAGGTCAAATTGCCAGCATCCAGCGAGGC 691

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
|||||
Db 692 ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGAGGAGTCTGTCTCGGGTCTTTGCT 751

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
|||||
Db 752 GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 811

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
|||||
Db 812 CAGCGGCAGTACATCTTCGAATACGATATGTGGACCGCCTGTCTGCCATCACCATGCC 871

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
:::|||||
Db 872 AGTGTGGCTCGCCACACCATGCAGACCATCCGATCCATTGGCTACTACCGCAACATATAC 931

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 932 AACCCCGGAAAGCAACGCTCCATCATCACGGACTCAACAGGAGGGCTGCTTCTA 991
QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 992 CAAACAGCTTCTTGGGTACAAGTCGGAGGCTCTTATTCAAATACAGAAGGCACAGTAGG 1051
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 1052 CTCTCAGAAATTTTATATGATAGCACAGAGTCAGTTTTTACCTATGATGAACACAGCAGGA 1111
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheArgPheThrCysThrIleArgTyrArgGln 300
Db 1112 GTCTTAAAGACAGTAAACCTCCAGAGTGATGGTTTTTATTTCACCATTAGATACAGGCAA 1171
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 1172 ATTGTCCTCCCTGATTGACAGGCAGATTTTCGCTTTTAGTGAAGATGGGATGGTAATGCA 1231
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 1232 AGATTGACTATAGCTATGACACAGAGCTTTTCGAGTGACCAGCATCGAGGGTGTGATCAAT 1291
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValThrSerMetGlnAlaValIleAsn 360
Db 1292 GAAACGCCACTGCCTATTGATCTGTATCAGTTTGTATGACATTTCTGGCAAAAGTTGAGCAG 1351
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 1352 TTTGGAAAGTTTGGAGTTATATATTATGATATTAAACAGATCAATTTCTACAGCTGTAATG 1411
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 1412 ACCTATACGAAGCACTTTGATGTCTCATGGCCGTATCAAGGAGATTCATATGAGATATTC 1471
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 1472 AGGTCGCTCATGTACTGGATTACAATTCAATGATGATAACATGGGTCGGGTAAACCAAGAGA 1531
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 1532 GAGATTAAATAAGGGCCCTTTTGCCAACACCAACCAATATGCTTATGAAATATGATGTTGAT 1591
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db 1592 GGACAGCTCCAAACAGTTTACCTCAATGAAAAAGATAATGTGGCGGTACAACTACGATCTG 1651
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 1652 AATGGAAACCTTCATTTACTGAACCCAAAGTAACAGTGGCGTCTGACACCCCTTCGCTAT 1711
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 1712 GACCTGCGAGACAGAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT 1771
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 1772 TTCCTACGTCAAAGGGCAGCGAAATCTTTGAATATAGCTCCAAGGGGCTTCTAACTCGA 1831
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 1832 GTTTACAGTAAAGGCGAGTGGCTGGACAGTGATCTACCGTTATGACGGCCTGGGAAGGCGT 1891
QY 541 Val 541
Db 1892 GTT 1894

RESULT 7

US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, Kimberly A

; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:

Pred. No.: 1.11e-256 Length: 8473
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-038-854-39 (1-8473)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5130 TACCAGATTGGTTATGACGGCTCCCTCAGATTATCTACGCCAGTGGCCTGCACAC 5189

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5190 TACCAACAGAGCCGACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAAAC 5249

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5250 ATGACTTTGCCGTGGCAGAACGGTCAAACACTTGGTGAATGGAGATTCCGAAAAGAGCAA 5309

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5310 GCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTT 5369

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5370 TCAGTTGACTTTGATCGAACAAACAAAGACAGAAAAGATCTATGACGACCACCGTAAATTT 5429

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 5430 CTACTGAGGATCGCCTACGACACGCTCTGGGCACCCGACTCTCTGGCTGCGCAAGCAGCAAG 5489

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5490 CTGATGGCCGTCAATGTCACTATTCAATCCACAGGTCAAATTGCCAGCATCCAGCGAGGC 5549

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5550 ACCACTAGCGAGAAAGTAGATTATGACGACAGGGGAGGATCGTGTCTCGGGTCTTTGGCT 5609

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5610 GATGGTAAACATGGAGTTACACATATTAGAAAAGTCCATGGTCTCTCGGTCTTTGGCT 5669

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5670 CAGCGGCAGTACATCTTCGAATACGATATGTGGGACCGCCTGTCTGCCATCACCATGCC 5729

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5730 AGTGTGGCTGCCACACCATGCAGACCATCCGATCCATTGGCTACTACCGCAACATATAC 5789

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 5790 AACCCCCCGGAAAGCAACGCCCTCATCATCACGGACTACACGAGGAAGGGCTGCTTCTA 5849

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 5850 CAAACAGCTTTCTTGGGTACAAGTCGGAGGCTCTTATCAAATACAGAGGCAGACTAGG 5909

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 5910 CTCTCAGAAATTTTATATGATAGCACAGACTCAGTTTACCTATGATGAAACAGCAGGA 5969

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 5970 GTCCTAAAGACAGATAAACCTCCAGAGTGATGGTTTATTGCAACATTAGATACAGGCAA 6029

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6030 ATTGGTCCCTGATGACAGGCAGATTTTCCGCTTAGTGAAGATGGGATGGTAAATGCA 6089

QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6090 AGATTGACTATAGCTATGACAACAGCTTTCCAGTGACCAGCATGCAGGGTGTGATCAAT 6149

QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db 6150 GAAACGCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAAGTTGAGCAG 6209

QY 361 pheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 6210 TTTGGAAAGTTTGGAGTTATATATTATGATATTAAACCAGATCATTTCTACAGCTGTAATG 6269

QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400

Db 6270 ACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC 6329

QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallyLys 420
Db 6330 AGTGTGCTCATGTACTGGATTACAATTCAATTCAGTATGATAACATGGGTGCGGTAAACCAAGAGA 6389

QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6390 GAGATTAAATAGGGCCCTTTGCCAACACCAACCAAAATATGCTTATGAATATGATGTTGAT 6449

QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db 6450 GGACAGTCCAAACAGTTTACCTCAATGAAAGATAATGTGGCGGTACAACTACGATCTG 6509

QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6510 AATGGAACCTCCATTACTGAACCCCAAGTAACAGTGGCGTCTGACACCCCTTCGCTAT 6569

QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6570 GACCTGCGAGACAGAAATCACTCGACTGGTGATGTTCAATATCGGTTGGATGAAGATGGT 6629

QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 6630 TTCCTYACGTCAAAGGGGCACGGAATCTTTGAATATAGTCTCAAGGGGCTTCTAACTCGA 6689

QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 6690 GTTTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCGTTATGACGGCCTGGGAAGCGGT 6749

QY 541 Val 541
Db 6750 GTT 6752

RESULT 8
US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29

```

; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 1,11e-256 Length: 8487
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-038-854-41 (1-8487)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5126 TACCAGATTGGTTATGACGGCTCCCTCAGAAATTATCTACGCCAGTGGCTGGACTCACAC 5185

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5186 TACCAACACAGAGCCGACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAAAC 5245

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5246 ATGACTTTTGCCTGGCGAGAACGGTCAAAACTTGGTGGATGGAGATTCCGAAAAGAGCAA 5305

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5306 GCCCAAGGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTT 5365

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5366 TCAGTTGACTTTGATCGAACAAACAAAGACAGAAAAGATCTATGACGACCACCGTAAATTT 5425

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArg 120
Db 5426 CTACTGAGGATCGCCTACGACACGCTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAG 5485

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5486 CTGATGCCCGTCAATGTCAACCTATTCATCCACAGGTCAAAATGCCAGCATCCAGCGAGGC 5545

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5546 ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTTGCT 5605

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5606 GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCCATGGTTCTTCTGCTTCATAGC 5665

```


QY	341	GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln	360
Db	6322	GAACGCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAGTGTGAGCAG	6381
QY	361	PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet	380
Db	6382	TTTGGAAAGTTTGGAGTTATATATTATGATTAACCAAGATCATTTCTACAGCTGTAATG	6441
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
Db	6442	ACCTATACGAAGCACTTTGATGCTCATGGCCGTTATCAAGGAGATTCAATATGAGATATTC	6501
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420
Db	6502	AGGTCGCTCATGTACTGGATTACAATTACAGTATGATAACATGGGTGCGGTAAACCAAGAGA	6561
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
Db	6562	GAGATTAAATAGGGCCCTTTGCCAACACCACCAATATGCTTATGAATATGATGTTGAT	6621
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu	460
Db	6622	GGACAGCTCCAAACAGTTTACCTCAATGAAAGATAATGTGGCGGTACAACACTACGATCTG	6681
QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
Db	6682	AATGGAAACCTCCATTTACTGAACCCCAAGTAACAGTGCCTCTGACACCCCTTCGCTAT	6741
QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
Db	6742	GACCTGCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATCGT	6801
QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
Db	6802	TTCTACGTCAAAGGGGCACGAAATCTTTGAATATAGCTCCAAGGGGCTTCTAACTCGA	6861
QY	521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
Db	6862	GTTTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCCGTTATGACGGCCTGGGAAGGCGT	6921
QY	541	Val	541
Db	6922	GTT	6924

US-10-029-020-14_COPY_1760_2300 (1-541) x US-10-038-854-35 (1-8675)

QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla	20
Db	5314	TACCAGATTGGTTATGACGGCTCCCTCAGAAATTATCTACGCCAGTGGCCTGGACTCACAC	5373
QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValIcIlyLysArgAsn	40
Db	5374	TACCAAAACAGAGCCGCACGTTCTGGCTGGCAACCGCTAATCCGACGGTTCGCCAAAAGAAAC	5433
QY	41	ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln	60
Db	5434	ATGACTTTGCCCTGGCGAGAACGGTCAAAACTTGGTGGAAATGGAGATTCCGAAAAGAGCAA	5493
QY	61	AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu	80
Db	5494	GCCCAAGGGAAAGTCAATGTCTTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTT	5553
QY	81	SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe	100
Db	5554	TCAGTTGACTTTGATCGAACAAACAGACAGAAAAGATCTATGACGACCCCGTAAATTT	5613
QY	101	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg	120
Db	5614	CTACTGAGGATCGCTACGACACGCTCTGGGCACCCGACTCTCTGGCTGCCAAGCAGCAAG	5673
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly	140
Db	5674	CTGATGGCCCGTCAATGTCACCTATTTCATCCACAGGTCAAATTGCCAGCATCCAGCGAGGC	5733

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5734 ACCACTAGCGAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTGCT 5793
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
Db 5794 GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 5853
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5854 CAGCGGCAGTACATCTTCGAATACGATATGTGGACCGCTGTCTGCCATCACCATGCC 5913
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5914 AGTGTGGCTCGCCACACCATCGAGACCATCCGATCCATTGGCTACTACCGCAACATATAC 5973
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 5974 AACCCCGGAAAGCAACGCCCTCCATCATCACGGACTACAACGAGGAGGGCTGCTTCTA 6033
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6034 CAAACAGCTTTCTTGGGTACAGTCGGAGGCTCTTATTCAATATACAGAAGGCAGACTAGG 6093
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6094 CTCTCAGAAAATTTTATATGATAGCACAGAGTCAGTTTACCTATGATGAAACAGCAGGA 6153
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6154 GTCTAAAGACAGTAAACCTCCAGATGATGGTTTTATTGACCATTAGATACAGGCAA 6213
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6214 ATTGTCCCTGATTGACAGGCAGATTTTCGCTTTAGTGAAGATGGGATGGTAAATGCA 6273
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6274 AGATTGACTATAGCTATGACACACAGCTTTCGAGTGACCAGCATGCAGGGTGTGATCAAT 6333
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 6334 GAAACGCCACTGCCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG 6393
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 6394 TTTGGAAAGTTTGGAGTTATATATTATGATATTAACAGATCAATTTCTACAGCTGTAATG 6453
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 6454 ACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC 6513
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 6514 AGGTCGCTCATGTACTGGATTACAATTCACTATGATAACATGGGTGGGTAAACCAAGAGA 6573
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6574 GAGATTAAATAGGGCCCTTTGCCAACACCCCAAATATGCTTATGATATGATGTTGAT 6633
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db 6634 GGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACTACGATCTG 6693
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6694 AATGGAAACCTCCATTACTGAAACCCAAAGTAAACAGTCGCGCTCTGACACCCCTTCGCTAT 6753
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6754 GACCTGCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT 6813

QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 6814 TTCCTACGTCAAAGGGGCACGGAATCTTTGAATATAGTCCAAGGGCTTCTAACTCGA 6873
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 6874 GTTTACAGTAAAGGCAGTGGCTGGACAGTGTATCTACCGTTATGACGGCCTGGGAAGCGT 6933
QY 541 Val 541
Db 6934 GTT 6936
RESULT 11
US-09-808-602-74
; Sequence 74, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John I
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-808-602-74
Alignment Scores:
Pred. No.: 1.19e-245 Length: 8797
Score: 2141.50 Matches: 382
Percent Similarity: 87.25% Conservative: 90
Best Local Similarity: 70.61% Mismatches: 68
Query Match: 75.27% Indels: 1
DB: 9 Gaps: 1
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-808-602-74 (1-8797)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5384 TACCAGCTCTGCAATAATGGAACCTCGGGTGATGTACGCCAACGGCATGGCTGTACGC 5443
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5444 TTCCACAGTGAGCCCCACGTCCTCGCAGCACCATCACCCCCACCATCGGGCGCTGCAAC 5503
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5504 ATCTCTCGCCCATGGAGAAATGGCCTGAACTCCATCGAGTGGCGCTGAGGAAGAACAG 5563
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5564 ATCAAGGCAAGATCACCATCTTTGGGAGGAAGCTTCGGGTCCACCGAAGGAATCTCTG 5623
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5624 TCCATTGATTATACCGAAATATCCGTACGAGAGAAGATCTACGATGACCACCGGAAATTC 5683
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120

Db 5684 ACCCTGAGGATCATCTATGACCAGGTGGCGCCCTTCTCTGTGGTCCCGAGCAGTGGG 5743

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140

Db 5744 CTGGCAGCGCTCAATGTCTCCTACTTCTCAATGGCGCTTGGCGCGCTCCAGCGAGG 5803

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 5804 GCCATGAGCGAGAGACAGACATTGACAAGCAAGCCGGATCGTGTCCCGCATGTTCCGC 5863

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180

Db 5864 GACGGGAAAGTCTGGAGTTATTCTCTATCTTGACAAGTCCATGGTCTCTTCTGCTACAGAG 5923

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

Db 5924 CAACGTCAGTACATATTTGAATATGACTCTCCGATCGCCTCCACGAGTCACTATGCC 5983

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 5984 AGTGTCCCGGCACAGCATGTCCACGCACACCTCCATTGGTTATACATCCGAAACATTTAC 6043

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 6044 AACCCACCCGAAAGCAATGCATCGGTCACTTTGACTACAGTACGCGCGCATCCTA 6103

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260

Db 6104 AAGACATCTTCTTGGGCACITGGCGCCAGGTGTTCTACAAGTATGGAAACTCTCCAAG 6163

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 6164 TTATCAGAGATAGTCTACGACAGCAGCAGCGCTCACCTTTGGGTATGACGACCCCGGT 6223

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db 6224 GTCCTGAAGATGGTCAATCTCCAAAGTGGGGCTTCTCCTGTACCATCAGTACCGAAAG 6283

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320

Db 6284 GTTGGGCCCTTGTGGACAAGCAGATTTACAGGTTCTCTGAGGAAGATGATCAACGCC 6343

QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339

Db 6344 AGGTTGATTATACCTATACGACAATAGCTTCCGCAATTGGCAGCATCAAAACCCGTCAAT 6403

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359

Db 6404 AGCGAGACTCCCCTTCCCTGTGACCTCTACCGCTATGACGAGATTTCCGGCAAGTGGAA 6463

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379

Db 6464 CACTTCGGCAAGTTTGGGTCATCTACTACGACATCAACAGATCATCACCACCTGCCGTC 6523

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399

Db 6524 ATGACGCTTAGCAAGCACTTTGACACCCATGGCGGCATCAAGGAAGTGCAATATGAGATG 6583

QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLys 419

Db 6584 TTCCGGTCCCTCATGTACTGGATGACTGTGCAATATGACAGTATGGGTAGGTATCAAG 6643

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439

Db 6644 AGGGAAC TGAACTAGGGCCCTATGCCAACACCAAAAGTACACCTATGACTATGACGGG 6703

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459

Db 6704 GACGGCCAGCTCCAGAGTGTGGCCGTCAATGACCGGCCTACCTGGCGCTATAGCTATGAC 6763

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479

Db 6764 CTCAATGGGAACCTGCACCTTCTAAACCCAGGAAACAGTGTCTCGCCTCATGCGCCTTACGC 6823

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499

Db 6824 TATGACCTCCGTGACCGGATAACCAGGCTAGGGGACGTGCAGTACAAAATCGATGACGAT 6883

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519

Db 6884 GGCTATTTGTGCCAGAGAGGGTTCAGACATCTTTGAATACAACCTCCAGAGGCCTTCTGACG 6943

QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539

Db 6944 AGAGCATACAACAAGCCAGCGGATGGAGCGTGACGTACCGCTATGACGAGTGGGCCGC 7003

QY 540 Arg 540

Db 7004 CGG 7006

RESULT 12

US-09-808-602-77

; Sequence 77, Application US/09808602

; Patent No. US20020155115A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Corine A

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrman, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishnu

; APPLICANT: Mezes, Peter S

; APPLICANT: MacDougall, John

; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nuclec Acids Encoding Same

; FILE REFERENCE: 15966-697 CIP

; CURRENT APPLICATION NUMBER: US/09/808,602

; CURRENT FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 09/800,198

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 77

; LENGTH: 8797

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-808-602-77

Alignment Scores:

Pred. No.: 1.19e-245 Length: 8797

Score: 2141.50 Matches: 382

Percent Similarity: 87.25% Conservative: 90

Best Local Similarity: 70.61% Mismatches: 68

Query Match: 75.27% Indels: 1

DB: 9 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-808-602-77 (1-8797)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db 5384 TACCAGCTCTGCAATAATGAACCTCGGGTGATGTACCCAACGGCATGGGCTGTGACG 5443

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 5444 TTCCACAGTGAGCCCCACGTCTCTCGCAGGCACCATCCCCCACCATCGGGCGGTGCAAC 5503

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60

Db 5504 ATCTCTCTGCCCATGGAGAATGGCCTGAACCTCCATCGAGTGGCGCTGAGGAAGAACAG 5563

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80

Db 5564 ATCAAAGGCAAAAGTCACCATCTTTGGGAGGAAGCTTCGGGTCCACGGAAGGAATCTCTG 5623

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100

Db 5624 TCCATTGATTATGACCGGAATATCCGTACGGAGAAGATCTACGATGACCACCGGAATTC 5683

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArg 120

Db 5684 ACCCTGAGGATCATCTATGACCAGGTGGCGCGCCCTTCCTGTGGCTCCCGAGCACTGGG 5743

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140

Db 5744 CTGGCAGCCGTCAATGTCTCCTACTTCTTCAATGGCGCGTGGCCGGCCTCCAGCGAGGG 5803

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 5804 GCCATGAGCGAGAGGACAGACATTGACAAGCAAGGCCCGGATCGTGTCCCGCATGTTGCC 5863

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180

Db 5864 GACGGGAAAGTCTGGAGTTATTCTCTATCTTTGACAAGTCCATGGTCTCTCTGCTACAGC 5923

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

Db 5924 CAACGTCAGTACATATTGAATATGACTCTCCGATCGCCTCCACGCAGTCACTATGCC 5983

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 5984 AGTGTGCCCCGGCACAGCATGTCCACGCACACCTCCATTGTTTACATCCGAAACATTAC 6043

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 6044 AACCCACCCGAAAGCAATGCATCGGTCACTCTTTGACTACAGTATGACGGCCGCATCCTA 6103

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleIleTyrLysTyrGlyLysLeuSerLys 260

Db 6104 AAGACATCTTTCTTGGGCACCTGGGCGCCAGGTGTTCTACAGTATGGAAACTCTCCAG 6163

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 6164 TTATCAGAGATAGTCTACGACAGCACAGCCGTCACTTTGGGTATGACGAGACCACCGT 6223

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db 6224 GTCTGAAGATGGTCAATCTCCAAAGTGGGGCTTCTCCTGTACCATCAGGTACCCGAAAG 6283

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320

Db 6284 GTTGGGCCCTTGTGGACAAGCAGATTACAGGTTCTCTGAGGAAGGAATGATCAACGCC 6343

QY 321 ArgPheAspTyrAsnTyr--AspAsnSerPheArgValThrSerMetGlnAlaValIle 339

Db 6344 AGGTTGATTATACCTATCACGACAATAGCTTCCGCATTCGCCAGCATCAAAACCCGTCA 6403

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359

Db 6404 AGCGAGACTCCCCTTCCTGTGTGACCTCTACCGCTATGACGAGATTTCGGCAAGGTGGA 6463

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379

Db 6464 CACTTCGGCAAGTTTGGGGTCACTACTACGACATCAACCAGATCATCACCACGTGCCGTC 6523

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399

Db 6524 ATGACGCTTAGCAAGCACTTTGACACCCATGGGCGCATCAAGGAAGTGAATATGAGATG 6583

QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLys 419

Db 6584 TTCCGGTCCCTCATGTACTGGATGACTGTGCAATATGACAGTATGGGTAGGTCAATCAAG 6643

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439

Db 6644 AGGGAAGTGAACACTAGGGCCCTATGCCAACACCAAGTACACCTATGACTATGACGGG 6703

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459

Db 6704 GACGGCCAGCTCCAGAGTGTGGCCGTCAATGACCGGCCCTACCTGGCGCTATAGCTATGAC 6763

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479

Db 6764 CTCAATGGGAACCTGACCTTCTAAACCCAGGAACAGTGTGCTCATGCCCTTACGC 6823

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499

Db 6824 TATGACCTCCGTGACCGGATAACCAAGGCTAGGGGACGTGACGTACAAATCGATGACGAT 6883

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519

Db 6884 GGCTATTGTGCCAGAGAGGTCAGACATCTTTGAATACAACTCCAAGGGCCCTCTGACG 6943

QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539

Db 6944 AGAGCATACAACAAGCCACGGATGGAGCGTGACGTACCGCTATGACGGAGTGGCGCG 7003

QY 540 Arg 540

Db 7004 CGG 7006

RESULT 13

US-09-800-198-62

; Sequence 62, Application US/09800198

; Publication No. US20030087816A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrmann, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 62

; LENGTH: 8797

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-800-198-62

Alignment Scores:

Pred. No.: 1.19e-245 Length: 8797

Score: 2141.50 Matches: 382

Percent Similarity: 87.25% Conservative: 90

Best Local Similarity: 70.61% Mismatches: 68

Query Match: 75.27% Indels: 1

DB: 10 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-800-198-62 (1-8797)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db 5384 TACCAGCTCTGCAATAATGGAACCCCTGCGGTGATGTACGCCAACGGCATGGCTGTGACG 5443

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 5444 TTCCACAGTGAGCCCCACGTCCTCGCAGGCACCATCACCCGCCCATCGGCGCTGCAAC 5503

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60

Db 5504 ATCTCTCTGCCCATGGAGATGGCCTGAACTCCATCGAGTGGCGCCTGAGGAAGGAACAG 5563

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80

Db 5564 ATCAAGGCAAGTCAACATCTTTGGGAGGAAGCTTCGGGTCCACGGAAGGAATCTCCTG 5623

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100

Db 5624 TCCATTGATTATGACCGAAATATCCGTACGGAGAAGATCTACGATGACCACCGGAAATTC 5683

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120

Db 5684 ACCCTGAGGATCATCTATGACCGGTGGCGGCCCTTCCTGTGGCTCCCGAGCAGTGGG 5743

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140

Db 5744 CTGGCAGCCGTCAATGTCTCCTACTTCTCAATGGCGGCTTGGCGGCCCTCCAGCGAGGG 5803

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 5804 GCCATGACGAGAGGACAGACATTGACAAGCAAGCCGGATCGTGTCCCGCATGTTCCGCC 5863

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180

Db 5864 GACGGGAAAGTCTGGAGTTATTCCCTATCTTGACAAGTCCATGGTCTTCTGTACAGAGC 5923

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

Db 5924 CAACGTCAGTACATATTGAATAGACTCCTCCGATCGCCTCACGCAGTCACATATGCC 5983

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 5984 AGTGTGCCCGGCACAGCATGTCCACGCACACCTCCATTGGTTACATCCGAAACATTAC 6043

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 6044 AACCCACCCGAAAGCAATGCATCGGTCACTTTGACTACAGTATGACGGCCGATCCTA 6103

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260

Db 6104 AAGACATCTTTCTTGGGCACTGGCGGCCAGGTGTTCTACAAGTATGGAAAACCTCTCCAAG 6163

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 6164 TTATCAGAGATAGTCTACGACAGACAGCCGTCACCTTTGGGTATGACGAGACCACCGGT 6223

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db 6224 GTCTGGAAGATGGTCAATCTCCAAAGTGGGGCTTCTCCTGTACCATCAGGTACCGGAAAG 6283

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320

Db 6284 GTTGGGCCCCCTTGTGGACAAGCAGATTACAGGTTCTCTGAGGAAGGAATGATCAACGCC 6343

QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339

Db 6344 AGGTTGATTATACCTATACGACAATAGCTTCCGCAATGCCAGCATCAAAACCCGTCATT 6403

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359

Db 6404 ACGGAGACTCCCCCTTCTGTGACCTCTACCGCTATGACGAGATTTCCGGCAAGGTGGA 6463

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379

Db 6464 CACTTCGGCAAGTTTGGGGTCACTACTACGACATCAACACAGATCATCACCATGCCGTC 6523

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399

Db 6524 ATGACGCTTAGCAAGCACTTTGACACCCCATGGCGGCATCAAGGAAGTGAATATGAGATG 6583

QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLys 419

Db 6584 TTCCGGTCCCTCATGTACTGGATGACTGTGCAATATGACAGTATGGTAGGTTCATCAAG 6643

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439

Db 6644 AGGGAACCTGAAACTAGGGCCCTATGCCAACACCAACAAAGTAGACCTATGACTATGACGGG 6703

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459

Db 6704 GACGGCCAGCTCCAGAGTGTGGCGTCAATGACCGGCCTACCTGGCGCTATAGCTATGAC 6763

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479

Db 6764 CTCAATGGGAACCTGCACCTTTAAACCCAGGAAACAGTGTCTGCCTCATGCCCTTACGC 6823

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499

Db 6824 TATGACCTCCGTGACCCGGATAACAGGCTAGGGACGTGCAGTACAAATCGATGACGAT 6883

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519

Db 6884 GGCATTATTGTGCCAGAGAGGGTCAGACATCTTTGAATACAACTCCAAGGGCCTTCTGACG 6943

QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539

Db 6944 AGAGCATACAAAGGCCAGCGGATGGAGCGTGCAGTACCGCTATGACGGAGTGGGCCGC 7003

QY 540 Arg 540

Db 7004 CGG 7006

RESULT 14

US-09-800-198-65

; Sequence 65, Application US/09800198

; Publication No. US20030087816A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrmann, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 65

; LENGTH: 8797

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-800-198-65

Alignment Scores:

Pred. No.: 1.19e-245 Length: 8797

Score: 2141.50 Matches: 382

Percent Similarity: 87.25% Conservative: 90

Best Local Similarity: 70.61% Mismatches: 68

Query Match: 75.27% Indels: 1

DB: 10 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-800-198-65 (1-8797)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db 5384 TACCAGCTCTGCAATATATGAACCCCTGCGGGTATGTACGCCAACCGCATGGCTGTGCAGC 5443

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 5444 TTCCACAGTGAGCCCCACGTCTCTCGCAGGCACCCATCACCCCGCTCGGCGCTGCAAC 5503

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60

Db 5504 ATCTCTCTGCCCCATGGAGAATGGCCTGAACTCCATCGAGTGGCGCCTGAGGAAGGAACAG 5563

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5564 ATCAAAGGCAAAGTCACCATCTTGGGAGGAAGCTTCGGGTCCACGGAAGAAATCTCCTG 5623
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5624 TCCATTGATTATGACCGGAAATATCCGTACGGAGAAGATCTACGATGACCACCGGAAATTC 5683
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 5684 ACCCTGAGGATCATCTATGACCGAGTGGCGCCCTTCTGTGGCTCCCGAGCAGTGGG 5743
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5744 CTGCAGCCGTCAATGTCTCCTACTTCTTCAATGGGCGCTTGGCGGCTCCAGCGAGGG 5803
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5804 GCCATGAGCGAGAGGACAGACATTGACAAGCAAGCCGGATCGTGTCCCGCATGTTCCGC 5863
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5864 GACGGGAAAAGTCTGGAGTTATTCCTATCTTGACAAGTCCATGGTCTTCTGTACAGAGC 5923
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5924 CAACGTCAGTACATATTGATATGACTCTCCGATCGCCTCCAGCAGTCACATATGCC 5983
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5984 AGTGTGCCCGGCACAGCATGTCCAGGCACACCTCCATTGGTTACATCCGAAACATTTAC 6043
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 6044 AACCCACCCGAAAGCAATGCATCGGTCACTTTTGACTACAGTGATGACGGCCGATCCTA 6103
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6104 AAGACATCTTTCTTGGCACTGGGCGCCAGGTGTTCTACAGTATGGAATACTCTCCAAG 6163
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6164 TTATCAGAGATAGTCTACGACAGCAGCAGCCGTCACCTTTGGGTATGACGAGACCACCGT 6223
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6224 GTCCTGAAGATGGTCAATCTCCAAAGTGGGGCTTCTCCTGTACCATCAGGTACCCGAAAG 6283
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6284 GTTGGGCCCCCTTGTGGACAAGAGAGATTACAGGTTCTCTGAGGAAGGAATGATCAACGCC 6343
QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db 6344 AGGTTTGATTATACCTATCACGACAATAGCTTCCGCAITTCGCAATGCCAGCATCAAAACCCGTCATT 6403
QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359
Db 6404 AGCGAGACTCCCTTCCTGTGACCTCTACCGCTATGACGAGATTTCCGGCAAGGTGGAA 6463
QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
Db 6464 CACTTCGGCAAGTTTGGGGTCATCTACTACGACATCAACCCAGATCATCACCACCTGCCGTC 6523
QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 6524 ATGACGCTTAGCAAGCACTTTTGACACCCCATGGGCGCATCAAGGAAGTGCATATAGATG 6583
QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallys 419
Db 6584 TTCCGCTCCCTCATGTACTGGATGACTGTGCAATATGACAGTATGGGTAGGTATCAAG 6643

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 6644 AGGGAAC TGAAACTAGGGCCCTATGCCAACACACCACAAAGTACACCTATGACTATGACGGG 6703
QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459
Db 6704 GACGCCAGCTCCAGAGTGTGGCGTCAATGACCGGCCTACCTGGCGCTATAGCTATGAC 6763
QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 6764 CTCAATGGGAACCTGCACCTTCTAAACCCAGGAAACAGTGTCTGCCTCATGCCCCTTACGC 6823
QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 6824 TATGACCTCCGTGACCGGATAACACAGGCTAGGGGACGTGCAGTACAAAATCGATGACGAT 6883
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 6884 GGCTATTGTGCCAGAGAGSGTCAACATCTTTGAATACAACTCCAAGGGCCTTCTGACG 6943
QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 6944 AGAGCATACAACAAGGCCAGCGGATGGAGCGTGCAGTACCGCTATGACGGAGTGGGCCGC 7003
QY 540 Arg 540
Db 7004 CGG 7006

RESULT 15
US-09-808-602-79
; Sequence 79, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-808-602-79

Alignment Scores:
Pred. No.: 2.55e-245 Length: 8409
Score: 2138.50 Matches: 384
Percent Similarity: 86.69% Conservative: 85
Best Local Similarity: 70.98% Mismatches: 71
Query Match: 75.17% Indels: 1
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-808-602-79 (1-8409)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValala 20
Db 5386 TACCAGCTCTGTAATAATAGGTACTTTGAGAGTGATGTATGCCAATGGCATGAGTATTAGC 5445
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 160.054 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
Perfect score: 2845
Sequence: 1 YYIGADGSLRLLLANGMEVA.....YNRAGSWSVRYRDGLGRRV 541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delopt 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat 06082004 112217 29301/app query.fasta 1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1 258 @runat 06082004 112217 29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	232.5	8.2	4881	4	US-09-252-991A-7234
2	208	7.3	4998	4	US-09-543-681A-304
3	207	7.3	4761	4	US-09-543-681A-1157
4	200	7.0	4134	4	US-09-711-164-296
5	194.5	6.8	3387	4	US-09-252-991A-14461
6	193.5	6.8	4791	4	US-09-328-352-1416
7	192	6.7	5519	4	US-09-453-702B-226
8	188.5	6.6	31960	4	US-09-453-702B-11
9	185.5	6.5	2415	4	US-09-328-352-1419
c 10	185.5	6.5	4453	4	US-09-453-702B-90
11	184	6.5	2997	4	US-09-252-991A-11853
12	182.5	6.4	2925	4	US-09-252-991A-7069

13	182	6.4	4884	4	US-09-328-352-2478	Sequence 2478, Ap
14	177.5	6.2	1857	4	US-09-252-991A-11787	Sequence 11787, A
15	176	6.2	1857	4	US-09-252-991A-7125	Sequence 7125, Ap
16	173.5	6.1	4320	4	US-09-543-681A-3388	Sequence 3388, Ap
c 17	169.5	6.0	11613	4	US-09-453-702B-42	Sequence 42, Appl
18	168.5	5.9	9717	3	US-09-251-645-1	Sequence 1, Appli
c 19	162.5	5.7	1683	4	US-09-252-991A-11588	Sequence 11588, A
c 20	161	5.7	1683	4	US-09-252-991A-6977	Sequence 6977, Ap
c 21	146.5	5.1	34094	4	US-09-292-034-1	Sequence 1, Appli
c 22	127	4.5	2181	4	US-09-252-991A-14909	Sequence 14909, A
23	125.5	4.4	1515	4	US-09-252-991A-11708	Sequence 11708, A
24	123	4.3	993	4	US-09-252-991A-14721	Sequence 14721, A
c 25	123	4.3	1764	4	US-09-252-991A-14908	Sequence 14908, A
c 26	121.5	4.3	14066	4	US-09-601-198-56	Sequence 56, Appl
27	120	4.2	3827	2	US-08-447-031A-1	Sequence 1, Appli
28	117.5	4.1	4320	4	US-09-134-000C-2728	Sequence 2728, Ap
c 29	116.5	4.1	1260	4	US-09-252-991A-15163	Sequence 15163, A
c 30	116.5	4.1	2241	4	US-09-252-991A-12657	Sequence 12657, A
31	115	4.0	1593	4	US-09-252-991A-7095	Sequence 7095, Ap
32	115	4.0	94750	4	US-09-596-002-38	Sequence 38, Appl
33	113	4.0	37948	3	US-09-251-645-11	Sequence 11, Appl
c 34	112.5	4.0	5903	4	US-09-453-702B-195	Sequence 195, App
35	112	3.9	627	4	US-09-252-991A-14601	Sequence 14601, A
36	112	3.9	4612	2	US-08-447-031A-8	Sequence 8, Appli
37	111	3.9	2412	4	US-09-252-991A-11898	Sequence 11898, A
c 38	111	3.9	4161	4	US-09-252-991A-12016	Sequence 12016, A
c 39	109.5	3.8	1664976	4	US-08-916-421B-1	Sequence 1, Appli
40	109	3.8	2651	1	US-08-462-949-1	Sequence 1, Appli
41	109	3.8	2651	1	US-08-023-764B-1	Sequence 1, Appli
42	108.5	3.8	9510	4	US-09-453-702B-256	Sequence 256, App
43	108	3.8	1722	4	US-09-134-001C-1976	Sequence 1976, Ap
44	107	3.8	8133	1	US-08-480-604A-5	Sequence 5, Appli
45	107	3.8	8133	2	US-08-405-496A-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-7234
; Sequence 7234, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7234
; LENGTH: 4881
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7234

Alignment Scores:			
Pred. No.:	5.15e-18	Length:	4881
Score:	232.50	Matches:	148
Percent Similarity:	32.17%	Conservative:	82
Best Local Similarity:	20.70%	Mismatches:	217
Query Match:	8.17%	Indels:	268
DB:	4	Gaps:	31
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-7234 (1-4881)			
QY	7	GlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAlaLeuGlnThrGluProHis	26
Db	2025	GGCCGGCGGCGCTCTTCTCTG---GGAATGGGAGGCGAGGCAA-----	2069

QY 27 LeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnValThrLeuProIleAsp 46
Db 2070 -----GCAGCCCGCGCGGTCCATCACTGGCCAG-----CTTCCCGCAGAT 2111
QY 47 Asn---GlyLeuAsnLeuValGluTrpArgGlnArg-----LysGluGlnAla 61
Db 2112 GGACTCGCGGTACGTCTGGAAACGAAGACGCAGCGTTACCGCGATCAACGCCGACGGCAG 2171
QY 62 ArgGlyGlnValThr-----ValPheGlyArgArgLeuArg 73
Db 2172 CGAGGAAGTCTATGTCCACGACGACAAACGCCGCGCTGGTACGCCAGGTCTGATCCCGACGG 2231
QY 74 ValHisAsnArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIle 93
Db 2232 AGGCGAAACCCCTCAGGCACCTACGACGAGAGAGGGCCAACTGGTCGCCGAACGCGATCCGTT 2291
QY 94 TyrAspAspHisArgLysPheThrLeuArgIle-LeuTyrAspGlnAlaGlyArgProse 113
Db 2292 GGGAACGATCAC-----CGAGTACCGCTACGACGAGAGCGGACGCTTGGA 2336
QY 113 rLeuTrpSerProSerArgLeuAsnGlyValAsnValThrTyrSerPro---GlyG1 132
Db 2337 GGCACGTCTCCCGCC-----GAAGCGAGCCAACTGCTATAGTACTTTCGACGG 2387
QY 132 yTyrIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaG1 152
Db 2388 CTTCGTGCGCAGCGTGCAGCGCGGAGGCGCAGTGGAAGTACGAGCGCAACGCTCAGGG 2447
QY 152 yArgIleThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLy 172
Db 2448 CGACATCACCCGCGACCGGATCCCGAAGGCAAGTCAACCCACTACGCCCTACGACCCCG 2507
QY 172 sSer---MetValLeuLeuHis----- 179
Db 2508 CGGCTGCCTGGTGGAATCACCCATGCCGACGGCCTGCACCCAGTTGACCTGGAATCC 2567
QY 180 -----SerGlnArgGlnTyr----- 184
Db 2568 GCTCGGCCAGTTGATAGAAGAACACAGCTTCCGACGGTAGCGTGCAGCGCTATCGTACGA 2627
QY 185 -----IlePheGluPh 188
Db 2628 CACACTGGGTAGACAGATCACCCGCCAGGACGAAAGCGGCGCATCACCCGGTTCCAATG 2687
QY 188 eAspLysAsnAspArgLeuSerSerValThrMetProAsnValAlaArgGlnThrLeuG1 208
Db 2688 GGATGCAGCAGCGCGCTGAGCCAGATCACCCCTGCCCGGTGGCGCCAGCCGCATG----- 2742
QY 208 uThrIleArgSerValGlyTyrTyrArgAsnIleTyr-----GlnProPr 223
Db 2743 -----TACCGCTACAACGCTTACGGCAAGGTCAACCGAGTGCGA 2783
QY 223 oGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPh 243
Db 2784 CGAACAGGCGCGATGACCCGCTACGAATATCTTGACGACCTGCACCTGGTCAGCCGGCG 2843
QY 243 eTyrLeuGlyThrGlyArgArgValIleTyrLysTyr----- 255
Db 2844 GATCAACCCGCGAGTCAGTTCAGTTCGCTACCGCTACGAAACGCCCGGCTACTGCTCAG 2903
QY 256 -----GlyLysLe 258
Db 2904 CGAGATCGAGAACGACGCGGGAACGCTATCGCCCTCGATTACCACGGCAACGGCCTGAT 2963
QY 258 uSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluTh 278
Db 2964 CAGCCAG-----GAAACCGGCTTCGACGCGCGCGCACCGCCTATCGCTACGACCTCAA 3017
QY 278 rAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTy 298
Db 3018 AGCCAGCTTCTGGAGAAACCGAGTACGGCGACGACGCGCAACTGCGCACCACTA 3077
QY 298 rArgGln-----IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluG1 316

Db 3078 CCAGCGGCACAGTACCGGGCGCCTGCTGGCGAAGACACTG-----CCCAGCGG 3125
QY 316 yMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetG1 336
Db 3126 C-----AATCGGGTCGACTACCGCTACGATACGCTCGGACGCGCTGGTCGCC----- 3171
QY 336 nAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerG1 356
Db 3172 -----GTCGACGACGCGCACCTGGCCGCTG---GCCTACGAATACGAC----- 3210
QY 356 yLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleTh 376
Db 3210 ----- 3210
QY 376 rThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValG1 396
Db 3211 -----CTACG 3215
QY 396 nTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGln-----TyrAspAsnMe 414
Db 3216 GGACCGCCTGGTCCGCGAGCACCGGGCTGGGCCACCTTCACTACGCTACGACGCCCT 3275
QY 414 tGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSe 434
Db 3276 CGGCCAGTTGATCCACTGCCGACTGCCCGACGGC-----AACCGCGTCGA 3320
QY 434 rTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTr 454
Db 3321 CTACCGCTACCAGACCGGCGGACCCCTCTCGGCAATCGATCTCAACCGCCAGTCCCTGAC 3380
QY 454 pArg----- 455
Db 3381 CCGGCACCAAGTTCGGTAGCGGTTCGCGAACCGCCAGCGCCAGGCGGAAGTCTCAGCCA 3440
QY 456 -TyrSerTyrAspLeuAsnGlyAsnLeu----- 465
Db 3441 ATACCACTACGATGAGCAGGCGCGCTGCTGGCCCCACCAGGTCACTCAGCGGCAGCGCCA 3500
QY 465 sLeuLeu-----SerProGlyAsnSerAlaArgLeuThrProLe 478
Db 3501 TCTCTACCAGCGCCAGTACCGCTATGACGTAGCGGCAACCTGGCCGCCATCGAGGACAG 3560
QY 478 uArgTyrAspIleArg-----AspArgIleThrArgLeuGlyAspValGln-- 493
Db 3561 CCGCAAGGGCATTCGACGCTTCCACTACGACCCGCTGGACCGCCTGCTCGGCGTACGCGG 3620
QY 494 -----TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyAs 508
Db 3621 TGAACCCCGGAAAGCTTCTGTGCACGACCCGCGCGCAACCTGTCTCGCCCGGCGGCA 3680
QY 508 pIle----- 509
Db 3681 ATTCGACGCGCGACAGATGGAGGTTTCGCGGAAATCGCCTGCTGACGCGAGGCGACCGGCA 3740
QY 510 -PheGluTyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGly----- 526
Db 3741 TTTCGACTACGACGCGCCACGCGCAATCTCGTTGCGGAACCGCGGCAACCGGCAAGCT 3800
QY 527 -----Se 527
Db 3801 GGTCAACCGAATACAGCTACGACTGCGCAACACCGCCTGATCGGGCTCAGCCTGCGGACGG 3860
QY 527 rTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArgVal 541
Db 3861 GCGCCAGGTCTGTACCGCTACGACGCCCTTCGCGCCGACGCATC 3903

RESULT 2

US-09-543-681A-304
; Sequence 304, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON


```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 304
; LENGTH: 4998
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; US-09-543-681A-304

Alignment Scores:
Pred. No.:      6.42e-15      Length:      4998
Score:          208.00      Matches:      140
Percent Similarity: 31.40%      Conservative: 87
Best Local Similarity: 19.36%      Mismatches:  224
Query Match:      7.31%      Indels:      272
DB:              4          Gaps:       37

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-543-681A-304 (1-4998)

QY      2 TyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAlaLeu 21
Db      2200 TACATTGACGCTGAGGCT-----GGATTTTCACGTTATTAT 2235

QY      22 GlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnVal 41
Db      2236 TATAATGATGATAATCTGGTGACTCGCACTATCGATCCA---TTATGGCGTGAACGCTAT 2292

QY      42 ThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGlnAla 61
Db      2293 ACG-----GAATGGGAACAGCGGAAAAAGATAGCA 2322

QY      62 Arg-----GlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArg 77
Db      2323 GAAATCAATGAGATTGGTGAGCGCACAGAGTATGGT-----TATCACTATAAT 2370

QY      78 AsnLeuLeuSerLeu-----Asp 83
Db      2371 GGACTGCTTGCCCTATATCTATTACCTGATGGTAAGCCATTTACTATGATTATAATGAC 2430

QY      84 PheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArg 103
Db      2431 TATGGTCAACTTACCCATTTCACATCGGCCCTTTGGCGACGAA-----TGGCAG 2478

QY      104 IleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArgLeuAsnGly 123
Db      2479 CTAAGCTATGATGAAATGGC-----AAT 2502

QY      124 ValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSer 143
Db      2503 TTAACCATAGTGACAGATCCGCAAGGG-----CGTCAG 2535

QY      144 GluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGlyLys 163
Db      2536 CAGGTCTATGATATATAGCCAAACATGGTGAATTGCTCAAAGCCATTACGCCCAATGGTGGC 2595

QY      164 ThrTrpSerTyrThrTyrLeuGluLysSerMetValLeu----- 176
Db      2596 CAATGGCAATACCACTATAACCCCTGCACATCAACTGATTAAACCACCAATCCTTATCAA 2655

QY      177 -----LeuLeuHis----- 179
Db      2656 CACAGCACAGATATCATAGTATGATGAATTGGGCGGTTTACTGCACCTACACTGATGCGTTA 2715

QY      180 -----SerGlnArgGlnTyrIlePheGluPhe---AspLysAsnAspArgLeuSer 195
Db      2716 AATCACACCACACGCTATCAATACAGCACTGAGCATGCCAGTACCACACGCGAGTCTCAGC 2775

QY      196 SerValThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyr 215
```

```
Db      2776 AAGATTATTATACCGAT-----GGGGTC 2799

QY      216 TyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGlu 235
Db      2800 GAGCAACAGATAGACTATGACAGTGAGCGCGGAGTTATCGCGGTGACGGAT----- 2850

QY      236 AspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyr 255
Db      2851 -----GGTGAAGGAAAAAACCCCGTTATCGCTTTGAA 2880

QY      256 GlyLysLeuSerLysLeuAlaGluThrLeuTyr---AspThrThrLysValSerPheThr 274
Db      2881 GGGCCTTTTGATGTCTTACTGGCAATGATACGCCCTGACGCGAGTGAATTCGCTTTGAA 2940

QY      275 TyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGlu----- 290
Db      2941 TATGATTCACTGACACGACTTAAAAAAGTAGTGAATCGGAACGGGAGGTTTATCTTTAC 3000

QY      291 -----GlyPheThrCysThr 295
Db      3001 GAACGGGATAAAGCGGGACAAATCATTCGTGAAGTTGATTTTACGGGCGTGAGATTGT 3060

QY      296 IleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIle----- 309
Db      3061 TATCGCTATGACCGACTAGGACGTGCGCATTCGCCACTCGCTATCCGATAACCATGAAGT 3120

QY      310 ---PheArgPheThrGluGluGlyMetVal----- 318
Db      3121 CGTTGGCGCTATAATGAAAGTGGCTTAGTGGTTGAACAAAGCGAATGGTTTGAGATGAG 3180

QY      319 -----AsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArg 331
Db      3181 CAAGAAAGCCGCTGTTTATCGACGGCTCAATATAGATATAAC----- 3222

QY      332 ValThrSerMetGlnAlaValIleAsnGluThr---ProLeuProIleAspLeuTyrArg 350
Db      3223 -----GCACGTCAACAATTGATAAAAGCGACTAATCCTGACTCGTGGTTGAAATTGAG 3276

QY      351 TyrAspAspValSerGlyLys-----ThrGluGlnPheGlyLysPheGlyValIleTyr 368
Db      3277 TATGACGAC---CAAGGCCGTTTATGCAGTGAACGTATTAAATGAACAAGAGATTGTT--- 3330

QY      369 TyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHis----- 385
Db      3331 CATCAATGGAATGAAGCTGATAATACGCTGGCATTGACTCGTTTTGGTGAGCGGAGTTG 3390

QY      386 -----PheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSer 402
Db      3391 CATTATGCTTTTGGGCGACTGGGGGAGTTAACTTCACTGCAAGTCAATCAGCATGCACCA 3450

QY      403 LeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeu 422
Db      3451 CTTCAATTC-----AGCTACAAACGCGGTAGGACAAGAGTATTATACGCCGTAGT 3498

QY      423 LysValGly----- 425
Db      3499 CGCGCCGGATTGTCAATTCCAGTCACTATACGGCTACCGGTTTACTGGCTCATCAACGA 3558

QY      426 -----Pro 426
Db      3559 GCAGGGCGAGGCACCGAACAGTTTTTTACAATCGATACAAGCCCATCCCCAACACGCCCA 3618

QY      427 TyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla----- 439
Db      3619 TTTTGTACGGATGTGCACCGCAGTTTATCAATATGATCGCGCTTATAATGTGGTGGGTATT 3678

QY      440 -----AspGlyGlnLeuGlnThr-----ValSerIleAsnAspLys----- 451
Db      3679 GAAGATGATCGCTGGCGACAGACACGCTTATCACTATAATGCCAATGACCAAAATCACTGAG 3738

QY      452 -----ProLeuTrp-----ArgTyrSerTyrAspLeuAsnGly 462
```

Db	3739	ACACAATACAGTCCACAATGGGGTAATCAGGACGAAAAAGTTCCAGTATGATAACAACCTC	3799
QY	463	AsnLeu-----HisLeuLeuSerProGlyAsnSer-----	472
Db	3799	AATATCAGGAGCACTTAACAACGCCGTCCTTCGTCGATGGTGCCTTCAGATGCACAAGGG	3858
QY	473	AlaArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArg-----	488
Db	3859	GCGATGCTACAGCTATTTCAGCAACAGCAAGCAGGACGGGTGACTCGCCGTTACACAGCG	3918
QY	489	LeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGln-----	504
Db	3919	AAGGGATATCAAGATTACCATTTATGATGTTAATGGCGGATTAGCTAAAAAATTGTCCAT	3978
QY	505	---ArgGlyGlyAspIlePheGlu-----TyrAsnSerAlaGlyLeuLeuIle	519
Db	3979	ACACGGGTTTTTCGTCCCAGAGAAATGGCGCTATTATTGGAACACACAAAAATCAGTGACC	4038
QY	520	LysAlaTyrAsnArgAlaGlySer---TrpSerValArgTyrArgTyrAspGlyLeuGly	538
Db	4039	GCATGCTTTACGCCCTAAAGGGGATTGTTGG-----CACTATACTTATGATGCTTTTGGT	4092
QY	539	ArgArgVal	541
Db	4093	CGACGTCGTG	4101

RESULT 3
 US-09-543-681A-1157
 ; Sequence 1157, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; CURRENT FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 1157
 ; LENGTH: 4761
 ; TYPE: DNA
 ; ORGANISM: Proteus mirabilis
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (4726)
 ; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
 US-09-543-681A-1157

Alignment Scores:	7.9e-15	Length:	4761
Pred. No.:	207.00	Matches:	148
Score:	32.66%	Conservative:	80
Percent Similarity:	21.20%	Mismatches:	218
Best Local Similarity:	7.28%	Indels:	253
Query Match:	4	Gaps:	33

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-543-681A-1157 (1-4761)

Qy	30	GlyThrValAsnProThrValGlyLysArgAsnValThrLeuProIleAspAsnGlyLeu	49
		:	:
Db	1793	GGGTCAGTAGTCATTATCGTTGGC-----AACGTTTTGATGCCTTTACCATTTGAAGAC-	1845
		:	:
Qy	50	AsnLeuValGluTrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPheGly	69
		:	:
Db	1846	AATGAACCTGAATGGCGAGTGGTTGAGCACTGGCTTAAAGAC-----GGT	1890
		:	:
Qy	70	ArgArgLeuArgValHisAsnArgAsnLeuSerLeuAspPheAspArgValThrArg	89
		:	:
Db	1891	AAACGCTGTTTAGAGCATACC-----GAACCTGACGTATGATTTAGCTCAACGA	1938
		:	:
Qy	90	ThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLeuTyrAspGlnAla	109
		:	:

```
Db 2872 CCCTTATCCCGCAGGTGACCACTTCGGTTACGATAAAGTGGGACGGTTGCTGTTTCGA 2931
QY 396 GlnTyrGluIlePheArgSer----- 402
Db 2932 GAAAAACGCCGATTATCGCACGGAATACCTTTATCAACCTTTGAGTGTGACGTTACGCCGA 2991
QY 403 -----LeuMetTyrTip----- 406
Db 2992 GTCCCATGGCGGTATGGCATGAAGCCGAACGCCGACCACCGCCCGGTAGACTAT 3051
QY 407 -----MetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLys 423
Db 3052 CAAGACGCCCTGACCTTTACCTACGATAAAGTGGGCGAGTTGGTTCGAGAAGCCAGTGCT 3111
QY 424 ValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeu 443
Db 3112 CGCGGTGAT-----TATCAGCATCACTATGATGTCTGGGTAAATATC 3153
QY 444 GlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeuAsnGlyAsn 463
Db 3154 ACTGCACCGAGTTGCCCATCAAAGGCGATTTGAATACCTGTATTACGGCTCTGGGCAT 3213
QY 464 LeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg----- 479
Db 3214 TTACAACAAACGCAATGCGCAGATAATGAGCAACTGACGGTATTGGCGGAATATCAACGA 3273
QY 480 -----Tyr 480
Db 3274 GACCGTTTACACCGAGAAACGTTGGCCACCAGTGGCGCCTTAGACAATGAACGGGCTAT 3333
QY 481 AspIleArgAspArgIleThrArg----- 488
Db 3334 GACTGTGGGGACGTTATTACGCACCAAGTGGCGGCCAAATGAATGCCTCACAGTTTGTG 3393
QY 489 -----LeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGly 506
Db 3394 ACCCGGTGATAGACCGTCGTTATCGTTGGGATAAAACGCAATCAACTTATCGAGCGCAGT 3453
QY 507 -----GlyAspIlePhe----- 510
Db 3454 GTCAGTTATGTCAAACGGCGGAGGTTTTTACCGCGGACATTGGTACTTACCACAGTTAT 3513
QY 511 GluTyrAsn----- 513
Db 3514 CAGTACGACCCCGTGGGCAATTGACGGCGCATTTAGGTGGTGCAACCGAACACTTT 3573
QY 514 -----SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAla----- 525
Db 3574 CTGTATGATGCAGCAGCCAATTTACTGACTCGTCGCACACCCGAAGCACCGCATAAACCA 3633
QY 526 -----GlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArgVal 541
Db 3634 GTACAAGGCGAGTGATAAATTTGATTATCGCTACGACGGTTTTTGGTCGCATGGTA 3687

RESULT 4
US-09-711-164-296
; Sequence 296, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERET
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711,164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 296
; LENGTH: 4134
; TYPE: DNA
```

```
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4134)
US-09-711-164-296

Alignment Scores:
Pred. No.: 4.72e-14 Length: 4134
Score: 200.00 Matches: 136
Percent Similarity: 35.48% Conservative: 79
Best Local Similarity: 22.44% Mismatches: 225
Query Match: 7.03% Indels: 166
DB: 4 Gaps: 31

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-711-164-296 (1-4134)
QY 4 GlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAlaLeuGlnThr 23
Db 718 GGTGCTGGCGTCACTTCGGCTGGTACTGACACGACGGCGCAGCGGGCAGAAGAAGCC 777
QY 24 GluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnValThrLeu 43
Db 778 CGGCAGCAGGCCATTTCGGCGGGACGGAACCGTCGCTTTTCTCTGAT-----ACCTG 831
QY 44 Pro-----IleAspAsnGlyLeuAsnLeu--ValGluTrpArgGln 56
Db 832 CCGGTTACACCGAATATGGCCGGACACCGCATCCGTCTGTCTGCCGTGGCTGACG 891
QY 57 ArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsn 76
Db 892 CACGACCCGGAA-----TACCCG 909
QY 77 ArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAsp 96
Db 910 GAGAAATTTACCTGCCGCGCGCTGGTGGC-----TATGGCTGG 948
QY 97 HisArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSer 116
Db 949 ACGCCACGGCGCAACTGGCGTGTGTATGACCGTAGTGGCAACACAGTGGCGC----- 1002
QY 117 ProSerSerArgLeuAsnGlyValAsnValThrTyrSerPro-----GlyGlyTyr 133
Db 1003 -----AGCTTTACTTACGATGATAAATACCGGGCCGGATG 1038
QY 134 IleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArg 153
Db 1039 GTGGCGCACCGTCACACGGCGCGCGGAAATCCGTACCGTTACGACACGACGGCGG 1098
QY 154 IleThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSer 173
Db 1099 GTGACAGAAACAGCTAAACCCGGCAGGCTTAAGCTACACGTATCAGTAT--GAGAAAGAC 1155
QY 174 MetValLeuLeu-----LeuHisSerGlnArgGlnTyr 184
Db 1156 CGCATCACCATCACCGACAGCCTGGACCGCGGTGAAGTGTGCACACGACGGCGAAGCC 1215
QY 185 IlePhe-----Gluphe 188
Db 1216 GGGCTGAAGCGGTGGTGAAAAAGGAACACGCGGACGCGTACGACGAGAGTCAGTTT 1275
QY 189 AspLysAsnAspArgLeuSerSerValThrMetProAsnValAlaArgGlnThrLeuGlu 208
Db 1276 GACGCCGTGGCAGGCTCAGGGCACAGACG-----GATGCCGACGACGACACAGAG 1329
QY 209 ThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSer 228
Db 1330 -----TACAGCCCGGATGTGGTGACGGGCCTC 1356
QY 229 ValIleGlnAspPheThrGluaspGlyHisLeuLeuHisThrPheTyr----- 244
Db 1357 ATCAGCGGCATAACCAACGCCCGGATGGCAGGGCATCGGCGTTTACTATATAACCACCAAC 1416
QY 245 -----LeuGlyThrGly-----ArgArgValIleTyrLysTyr 255
```


Db 1417 CAGTTAAGTCAGCCACCGGCGCTGAGGGCTGGAATTGCGCGGGAATATGATGAATTG 1476
QY 256 GlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyr 275
Db 1477 GGCGTCTG-----ATTGAGAAACTGCCCTGACGGCGGATATCACCCCGCTACCGTTAT 1530
QY 276 AspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThr 295
Db 1531 GATAATCCACACAGTGACTTACCCTGCGCAACGGAAGATGCCACCGGAGCGGAAACC 1590
QY 296 IleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGlu 315
Db 1591 ATGACGTGGAGCGGTTACGGT-----CAGTTGCTGAGCTTACCGACTGT 1635
QY 316 GlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMet 335
Db 1636 TCCGGTTATGTAACCGGTTATGACCATGAC-----CGCTTCGGGCAGATG 1680
QY 336 GlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArg---TyrAsp----- 352
Db 1681 ACGGCGGTGACCGCGAGGAAGGGCTG-----AGTCAGTACCGCGCATACGACAGCCGT 1734
QY 353 -----AspValSerGlyLysThrGluGlnPheGlyLysPheGly 365
Db 1735 GGACAGTTAATTGCCGTGAAAGACACGCGAGGCGCATGAAACGCGGTAT---GAATACAAC 1791
QY 366 ValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHis 385
Db 1792 ---ATCGCCGGTGACCTGACCGCGGTCTATTCGCCCGGACGCGAGCAAGAAACGGACACAG 1848
QY 386 PheAspAlaTyrGlyArg---MetLysGluValGlnTyrGluIlePheArgSerLeuMet 404
Db 1849 TACGATGCGTGGGGAAGCCCGTCCGTACCCAGCGCGGGCTAACGCGCAGTATG--- 1905
QY 405 TyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysVal 424
Db 1906 -----GAATACGATGCTGCCGGACGCGGTCTATCCGCCTGACCCAGTGAAAC 1950
QY 425 GlyProTyrAlaAsnThrThrArgTyrSer-----TyrGluTyrAsp 438
Db 1951 GGCAGCCAC---ACCACCTTCGTTACGATGTACTTGACCGGCTGATACAGGAACCGGC 2007
QY 439 AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyr 458
Db 2008 TTTGACGGCGCACACAG-----CGTTATCACCCAC 2037
QY 459 AspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeu 478
Db 2038 GACCTGACCGGCAAA-----CTTATCCGACGAGGATGAGGGTCTGTCAACCCACTGG 2091
QY 479 ArgTyrAspIleArgAspArgIleThrArg-----LeuGlyAspVal-----Gln 493
Db 2092 CACTATGACGAAGCAGACCGCCTCACGCCACCGCAGCGGTGAAGGTTGAAACCGCAGAGCGG 2151
QY 494 TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn 513
Db 2152 TGGCAGTATGACGAACGTTGGCTGGCTGACAGAC----- 2184
QY 514 SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArg 533
Db 2185 -----ATCAGCATATACGGAAGGACCGGTCAGGGTGGCGGTGCATTACAGG 2229
QY 534 TyrAspGlyLeuGlyArg 539
Db 2230 TATGATGAGAAAGCGCGG 2247

RESULT 5

US-09-252-991A-14461
; Sequence 14461, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14461
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14461

Alignment Scores:
Pred. No.: 1.66e-13 Length: 3387
Score: 194.50 Matches: 141
Percent Similarity: 32.62% Conservative: 87
Best Local Similarity: 20.17% Mismatches: 248
Query Match: 6.84% Indels: 223
DB: 4 Gaps: 28

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-14461 (1-3387)
QY 54 TrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArgArgLeuArg 73
Db 880 TGGCGGAGACATGCTGGTGCCTCCACGCGGCGGGTCTGTGAAGCAGATCGAGCGGGCGGC 933
QY 74 Val-----HisAsnArgAsnLeuLeuSerLeuAspPheAspArg 86
Db 934 TACGAGTGGGATGTCCATGTCTCCGACGCGGCGGGTCTGTGAAGCAGATCGAGCGGGCGGC 993
QY 87 ValThrArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgIle----- 104
Db 994 CTGACCCGTACCTTCCGTACCTGCGCGATGCCACCGAGGTGACGACAGCCTCGGCGC 1053
QY 105 -----LeuTyrAspGlnAlaGlyArg-----ProSerLeuTrpSerProSerSerArg 120
Db 1054 GTCGAACGCTACGAGTTCGCGGCGGAGGGTGGGACGCGCGCTGGACCGCGCTGCTCAGG 1113
QY 121 LeuAsnGlyValAsnValThrTyr----- 128
Db 1114 GCCGACGGCAGCCGTAGCGAGTTTCGACTACGACCTGTTCCGCCCGCTGTCGCCATGCCG 1173
QY 129 SerProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyr 148
Db 1174 GATCCGCTGGGACGGGAGACGCGCGCGCGCGCGATGGCCAGGGCGGATGCTCGAGGAG 1233
QY 149 AspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThr 168
Db 1234 GAGTCGCGGGCAAGCGCGCTATCGCAAGCGCGTGCACGAAGAGACCGGGCTGCTGGTC 1293
QY 169 TyrLeuGluLysSerMetValLeuLeuLeuHisSerGlnArgGlnTyrIlePheGluPhe 188
Db 1294 GAGCTGGAGGACCGCATG-----CAGCGCGCTGGACCTTCGAGCGC 1335
QY 189 AspLysAsnAspArgLeuSerSerValThrMetPro---AsnValAlaArgGlnThrLeu 207
Db 1336 GACGAACGGGCAATGCGACGACGGTGC CGCGCGCGCGCGCAGCACCCGCTACGCCCTAC 1395
QY 208 GluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAla 227
Db 1396 GAAGACCGCGCCTGCCCGACCGCGCGCGCGAATCTGTCGACCGCGCGCGGCGGAGCGG 1455
QY 228 SerVal----- 229
Db 1456 CGCCTGGAGTGGAACCGCTTCGGCCTGTTGGCCGCTTGACCGCATGTTCCGCCCGAGGTC 1515
QY 230 ---IleGlnAspPheThrGluAspGlyHisLeuLeuHisThr----- 242
Db 1516 TGGCGTTACGACTACGACAACGAAGACGTTTGGTGGCCAGCAGCAGCCCGCTGGGACAA 1575

QY 243 -----PheTyrLeuGlyThrGly 248
Db 1576 CTGACCCGGCGCGCTACGACCGCTGGGGAACTGATCGGCCTGGAATTGGCCGACGGC 1635
QY 249 ArgArgValIleTyrLysTyr-----GlyLysLeuSerLysLeuAlaGlu----- 263
Db 1636 AGCGCCTGAGCTATGAATACGAGCCCTTGGCGGCAGACCCGTATCGCCGATGCCGAG 1695
QY 264 -----ThrLeuTyrAsp----- 267
Db 1696 GGGACGCCACGCTGTTTCTGGGGGCACGGCGATCTGCTGGCGGGTTTCCGACGCC 1755
QY 268 ---ThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsn 286
Db 1756 GCGCGCGCGAATTGTCCTACTGCACAGCAA--GCCGCGAGGCTGGTCGCCCTGACC 1812
QY 287 LeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAsp 306
Db 1813 AACGAGAAC--GGCGTCCAGCGCAGTTCCGCTACGACCTGCTGGACCGCTCGGTGGAA 1869
QY 307 -----ArgGlnIlePheArgPheThrGluGluGlyMetValAsn 319
Db 1870 GAGACCGGCTTCGACGGTCTGCGCAACGCTATCGCTACAACGCCGCTGACGAACTGATC 1929
QY 320 AlaArgPheAsp-----TyrAsnTyrAspAsnSerPheArgVal 332
Db 1930 GCCCGCGAGGATGCCGAGCTCGCGAGACCACCTATGCTACGACCGCGATGGCGCGCTG 1989
QY 333 ThrSerMetGlnAlaValIleAsnGlu---ThrProLeuProIleAspLeuTyrArg--- 350
Db 1990 GCGAGCATCCGGGTTCCCGCCACCGACGACGCCGCCCTGGTCGAGCGTTACCGCTGG 2049
QY 351 -----Tyr 351
Db 2050 TTGGCCGACGGCGCTGGCCAGCGCCGGGGCGGATTCGAGGTGGCTACAGGTAC 2109
QY 352 AspAspValSerGly-----LysThrGluGlnPheGlyLysPheGlyValIleTyrTyr 369
Db 2110 GACGAGTTCGTTAACCTTGGCCTGGAAAGCCAGGTACATGCCGCGTTGGTCTACAGC 2169
QY 370 AspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHisPheAspAla--- 388
Db 2170 GTCGAGCACAGCACGACGCGCTAGGCGTGCGGCAGACCGCGCTACGGCGACGCGCG 2229
QY 389 -----TyrGly----- 390
Db 2230 CCGTGCCTGGTGACCTATGGCCCGGGTCACTGTCACGGCGCGTGGTTCGGTGGTGC 2289
QY 391 -----ArgMetLysGluValGlnTyrGluIlePheArg--- 401
Db 2290 GAGTGGCCTTCGAGCGCGATGCCCTGCACCGTGAGGTGCGTCGCGACGCCCGCGGAC 2349
QY 402 -----SerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVal 418
Db 2350 GGGCAGGATGACCGGTGTTTACCCAGGAGCGCCAGCACGGCCGCTGGGACGCCCTGCAG 2409
QY 419 LysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAsp 438
Db 2410 CGCAGCCGCTGCGTCTGGCCGGCGGTTCGACTGGCAGCGC--GGCTATCGCTACGAC 2466
QY 439 AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyr 458
Db 2467 GGACTTGGACAACCTGGTCGGCATCGACGACAACAGTATCCGTCGCTGCGCTACGAATAC 2526
QY 459 AspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeu 478
Db 2527 GACCTCGCGGTGCGCTG---CTGGCCAGCCGGCGGCTGGCGCGCGCCAGCACCTAC 2583
QY 479 ArgTyrAsp----- 481
Db 2584 CGATACGACGCGCGCGCAACCGCCTGGAGGGCGCTCGCGGAGTACCGCGCGGAGGACGCT 2643

QY 482 -----IleArg 483
Db 2644 CGCAGGCATTTCGCCGAGAACGAACTGTACCGTCCGGTTTCAGTCGTCGAGGTGCGC 2703
QY 484 -----AspArgIleThrArgLeuGly 490
Db 2704 GCGAACCCAGGCCGCGAAGGCCCGCGCGCTGGCGCGCAACCGCTGGAGAGGATCGCC 2763
QY 491 AspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyAspIlePhe 510
Db 2764 GCGAACCGCTATCGCTTCGATCGCTGGGCAACCTGGTCGAGCGTATCGGGCGCGACGGC 2823
QY 511 Glu-----TyrAsnSerAlaGlyLeuLeuIle-----LysAlaTyr 522
Db 2824 GAGCGCCTGCGCCTGGCCTACGACGGCGCGAGCGCTTGTCCATCTGACGCGCGATTAC 2883
QY 523 AsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArgVal 541
Db 2884 GCCGACGGCACGGCGCTGGAGCGCGCTACCGCTACGACGCGCTGTCCGCGACGGATC 2940
RESULT 6
US-09-328-352-1416
; Sequence 1416, Application US/09328352
; Patent No. 5562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1416
; LENGTH: 4791
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1416
Alignment Scores:
pred. No.: 3.96e-13 Length: 4791
Score: 193.50 Matches: 128
Percent Similarity: 30.43% Conservative: 89
Best Local Similarity: 17.95% Mismatches: 211
Query Match: 6.80% Indels: 285
DB: 4 Gaps: 29
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-328-352-1416 (1-4791)
QY 1 TyrTyrIleGlyAlaAspGly---SerLeuArgLeuLeuAlaAsnGlyMetGluVal 19
Db 1837 TATTACTTCGATCTTGATGGGTTTACTTATCGAATCGTTTAGCGGATGGCGTGAGTCA 1896
QY 20 AlaLeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArg 39
Db 1897 TGGTATAGCCGTGAT---AsnGly-----GGGAAAAAG 1920
QY 40 AsnValThrLeuProIleAsp----- 46
Db 1921 CGCATTACTCGACAAATCGATTTTGATGGCGGAGAGACTCAACAAGAATAACAATGATCAA 1980
QY 47 ----- 48
Db 1981 GATCAACTGTATAAATTGTTTCAGCCAAACGGTGGGATAATTTCGTTTTCATATAATAAG 2040
QY 49 ---LeuAsnLeuValGlu-----TrpArgGlnArgLysGlu 59
Db 2041 CAGGTAACCTTAGTTGAAATTAAAGACCCAGAGGTAGTATTTCGAAACGAGAATACGAT 2100
QY 60 GlnAlaArg-----GlyGlnValThrValPheGlyArg 70
Db 2101 GAAATAGAAATGTTAGTAAGAAAAATCAATCCATTAGGGCATATCACACAGTAC----- 2154
QY 71 ArgLeuArgValHisAsnArgAsnLeuLeuSerLeuAspPheArgValThrArgThr 90

Db	2155	-----AAATATAACAATGATAATCAACTGGTTGAAGTCATTGATGCAAAAGGTGGTGT	2208
QY	91	GluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLeuTyr---AspGlnAla	109
Db	2209	AAGAAAATT-----CAGTATAATGAATTGGGCGAGATGATTTCATATACCGATTGCTCA	2262
QY	110	GlyArgProSerLeuTrp-----SerProSerSer	119
Db	2263	GGTAAGAGTAGTACATGGGAATATGATGAAGATGGTGCACAGCTGAACAAATGCA	2322
QY	120	ArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArg	139
Db	2323	AATAATAAGTTGTACAGTATTTTACAGTACCACAAAGGCCGAGATAAAGACAGCTGCAA	2382
QY	140	GlyIle-----MetSerGluArgMetGluTyrAspGlnAlaGlyArgIle	154
Db	2383	AGTATTATTATCTCTGATGGTTAAAGAAATACTTTGAACATGATGAAGACGCTTA	2442
QY	155	ThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMet	174
Db	2443	CTTAACACATACCGATACCAAGGTTTAGTAACGGAATATAAGTACAATCAAGTTGGCTTA	2502
QY	175	Val-----	175
Db	2503	TTAGAGCAGCGTATTGATGCCAATCGACATAGTGTAGCTTATCAATGGGACAAACAAGGT	2562
QY	176	-----LeuLeuLeuHisSerGlnArgGlnTyrIlePhe	186
Db	2563	CGCATTCAAAGCTAATCAACACAGAAATCAAGCCGAATATTGTTGGATACAATCCGTAT	2622
QY	187	-----GluPheAspLys	190
Db	2623	GGATATTGTACGTGAACAGGCCTTTGTATGGAGAGAGAAAAGCACTATAGTTACAATGAA	2682
QY	191	AsnAspArgLeuSerSerValThrMetProAsnValAlaArgGln-----	205
Db	2683	AATGGTCGATTATTTCAAATCCGTCGACCAATATACTAGACCAATTTTGACTATTATGCA	2742
QY	205	-----	205
Db	2743	GATGGACAAATTGCTTCGAAAAGTTTACCACCTTACATACAGGGCAGAAACAACAGAG	2802
QY	206	-----ThrLeuGluThrIleArgSerVal	213
Db	2803	CAATTTGACTATAACCTCAATAGCCAACTGAGCCGAGCCAGTAATGAGTTAGCCAAATT	2862
QY	214	GlyTyrTyrArgAsnIle-----TyrGlnProPro	223
Db	2863	GATTATATCGCAATGCACTTGGCAGTTGGTTAGAGAGCATCAGCATTATAAAATTCCA	2922
QY	224	Glu-----GlyAsnAlaSer	228
Db	2923	GAACTGAAACCGCTCACAGCAGTATTACATTATGAATATGATGAGCTTGGAAAT-----	2976
QY	229	ValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGly	248
Db	2977	TTAATTAAACAATTTCGACCTGATGGACATACACTGAATCATCTTGTATTGTTTCAGGG	3036
QY	249	ArgArgValIleTyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThr	268
Db	3037	CAT-----ATTATGCAATTGGT-----TTAATAACCAA	3066
QY	269	ThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGln	288
Db	3067	GAGGTGTTTCATTTCAGCGTGATGATTG-----CATAGAGAAACAACACGATTACTA	3120
QY	289	AsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGln	308
Db	3121	GCAATGGGCTCATGCAAAACCAACAATATAATGATGTGGG---TTATTGAGTTCACAG	3177
QY	309	IlePheArgPheThrGluGluGly-----MetValAsnAlaArgPheAspTyr	324

Db	3178	TTCAATCAACCTGAGCAAGAAACTCAAGACTATTTTACAATATCAAGCACATAGAAAAATAT	3237
QY	325	AsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluThrProLeu	344
Db	3238	CACACGATAAAAACTATTTGCTTAGTCAAGTTGAA-----	3273
QY	345	ProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGlnPheGlyLysPhe	364
Db	3274	-----GATAGCCGTTTGGGTAAACTC	3294
QY	365	GlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLys	384
Db	3295	AATTATCAATATGACCAATTGGCGGTTTAATTGCAGCTCAAAAGTTTACATAAAACAGAA	3354
QY	385	HisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMet	404
Db	3355	AGCTTTAAT-----	3363
QY	405	TyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysVal	424
Db	3364	-----TTTGACCCCGCAGGTAACTTAATTGATTCTGAGTCGGTATTA	3405
QY	425	GlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGln	444
Db	3406	AGTCCT-----GCACAAATTAAA	3423
QY	445	ThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeuAsnGlyAsnLeu	464
Db	3424	AATAACCTTTATTAAGAGTTATAAAGGTAAACACTACCAATATGACGTTCAAGGTAATGTG	3483
QY	465	---HisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArg	483
Db	3484	ACAGAGATTATTCAAGCGGGCAAAACCTTAAACTGACT-----TGGGATAATCAA	3534
QY	484	AspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArg	503
Db	3535	AATAGACTTATTCGA-----AGTGATAATAATGGTCTAGTTACA	3573
QY	504	GlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsn	523
Db	3574	GAATATGGTTACGATGCTCTTGGTCGT-----CGACTCTACAAA	3612
QY	524	ArgAlaGlySerTrpSerValArgTyrArgTyrArgTyrAspGly	536
Db	3613	AAACAGCCAAAGATAACCTTATTGTTGGTGGGATGGA	3651

RESULT 7

US-09-453-702B-226

; Sequence 226, Application US/09453702B

; Patent No. 6365723

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Valerie

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/453,702B

; FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 226:
SEQUENCE CHARACTERISTICS:
LENGTH: 5519
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 226:
US-09-453-702B-226

Alignment Scores:
Pred. No.: 7.76e-13 Length: 5519
Score: 192.00 Matches: 133
Percent Similarity: 34.70% Conservative: 78
Best Local Similarity: 21.88% Mismatches: 227
Query Match: 6.75% Indels: 170
DB: 4 Gaps: 29

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-453-702B-226 (1-5519)

QY 4 GlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAlaLeuGlnThr 23
Db 822 GGTGCGGGCGTCACTTCGGCTGGTACTGACCACGAGCGCGCAGCGGGCAGAGAAGCC 881
QY 24 GluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnValThrLeu 43
Db 882 CGGCAGCAGGCCATTTCGGCGGGACGGAAACCGTCGGCTTTTCCTGAT-----ACCCTG 935
QY 44 Pro-----IleAspAsnGlyLeuAsnLeu---ValGluTrpArgGln 56
Db 936 CCGGGTTACACCGAATATGGCCGGGACAAACGGCATCCGTCTGTCTGCCGTGTGGCTGACG 995
QY 57 ArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsn 76
Db 996 CACGACCCGGAA-----TACCCG 1013
QY 77 ArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAsp 96
Db 1014 GAGAATTTACCTGCGCGCGCGTGTGCGC-----TATGGCTGG 1052
QY 97 HisArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSer 116
Db 1053 ACGCGCGCGGCGAACTGGCGGTGGTGTATGACCGTAGTGGCAAACAGGTGCGC-----1106
QY 117 ProSerSerArgLeuAsnGlyValAsnValThrTyrSerPro-----GlyGlyTyr 133
Db 1107 -----AGCTTTACTACGATGATAAATACCGGGCCGGATG 1142
QY 134 IleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArg 153
Db 1143 GTGGCGCACCGTCACACGGGCGCGCGGAATCCGTTACCGTTACGACAGCAGCGGCGG 1202
QY 154 IleThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSer 173
Db 1203 GTGACAGAACAGCTAAACCCGGCAGGCTTAAGCTACACGTATCAGTAT---GAGAAAGAC 1259
QY 174 MetValLeuLeu-----LeuHisSerGlnArgGln--- 183
Db 1260 CGCATCACCATCACCGACAGCCTGAACCGCGGTGAAGTCTCTGCACACGACGAGGTGAAGGC 1319
QY 184 -----TyrIlePheGluPhe 188
:::|

Db 1320 GGGCTGAAGCGGGTGGTGAATAAGAAACACGCGGACGCGCAGCGTCACGACAGTCAGTTT 1379
QY 189 AspLysAsnAspArgLeuSerSerValThrMetProAsnValAlaArgGlnThrLeuGlu 208
Db 1380 GACGCGGTGGCAGGCTCAGGGCACAGACG-----GATGCGCGCAGGCAGGACAAACAGAA 1433
QY 209 ThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSer 228
Db 1434 -----TACAGCCCGGATGTGGTGACGGGCCTC 1460
QY 229 ValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyr----- 244
Db 1461 ATCACGCGCATCACACGCGGATGGCAGGGCATCGGCGTTTACTATAACCACCACACAGC 1520
QY 245 -----LeuGlyThrGly-----ArgArgValIleTyrLysTyr 255
Db 1521 CAGTTAACGTCAGCCACCGGCCTGACGGGCTGGAATAACGCCGGAATATGATGAATGG 1580
QY 256 GlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyr 275
Db 1581 GGCGCTCTG-----ATTCAGGAATACTGCCCTGACGGCGATATCACCCGCTACCGTTAT 1634
QY 276 AspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThr 295
Db 1635 GATAATCCACACAGTGACTTACCTCGCAACGGAAGATGCCACCGGCAGCGGAAACC 1694
QY 296 IleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGlu 315
Db 1695 ATGACGTGGAGCCGTACGGT-----CAGTTGCTGAGCTTCACCGACTGT 1739
QY 316 GlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMet 335
Db 1740 TCCGGTTATGTAAACCGTTATGACCATGAC-----CGCTTCGGGCAGATG 1784
QY 336 GlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSer 355
Db 1785 ACGCGGTGCACCGCAGGAAGGGCTG-----AGTCAGTACCGCGCATACGACAGCCGT 1838
QY 356 GlyLys-----ThrGluGlnPheGlyLys 363
Db 1839 GGACAGTTAATTGCCGTGAAAGACACGCGGGCCATGAAACGCGGTATGAATACAAACGCC 1898
QY 364 PheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThr 383
Db 1899 GCCGGT-----GACCTGACCACCGTCATTCGCCGACGCGCAGCAGAAACGGG 1946
QY 384 LysHisPheAspAlaTyrGlyArg---MetLysGluValGlnTyrGluIlePheArgSer 402
Db 1947 ACACAGTACGATGCGTGGGAAAGCCATCTGTACCACGAGGCGGTCTGACGCGCAGT 2006
QY 403 LeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeu 422
Db 2007 ATG-----GAATACGATGCTGCCGACGCGGTCTATCCGCTGACCAGT 2048
QY 423 LysValGlyProTyrAlaAsnThrThrArgTyrSer-----TyrGlu 436
Db 2049 GAAAACGGCAGCCAC---ACCACCTTCGTTACGATGTACTCGACCGCGTGATACAGGAA 2105
QY 437 TyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyr 456
Db 2106 ACCGGCTTTGACGCGCGCACACAG-----CGTTAT 2135
QY 457 SerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThr 476
Db 2136 CACCACGACCTGACCGGCAAA-----CTTATCCGACGAGGATGAGGGGTGGTCACC 2189
QY 477 ProLeuArgTyrAspIleArgAspArgIleThrArg-----LeuGlyAspVal--- 492
Db 2190 CACTGGCACTATGACGAAGCAGACCGGCTCACGCACCGCAGCGGTGAACCGCA 2249
QY 493 ---GlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGlu 511
Db 2250 GAGCGCTGCAGTAGTACGAAACGCGGCTGGCTGACAGAC----- 2288
:::|

QY 512 TyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArg 531
||| ||| ::::
db 2289 -----ATCAGCCATATCAGCGAAGGGCACCGGGTGACGGTGCAAT 2327

Qy 532 TyrArgTyrAspGlyLeuGlyArg 539
pb 2328 TACGGGTATGATGACAAAGGCCGG 2351

RESULT 8

US-09-453-702B-11
; Sequence 11, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
;

TITLE OF INVENTION: No. 6365723e1 Sequences of E. coli O157
 NUMBER OF SEQUENCES: 265

ADDRESS: 1 South Pinckney Street
CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

FILE, 33701 ELIS
COMPIER READABLE FORM:

COMPUTER READABLE FORM.

MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453.702B

ETD.ING DATE: 03-Dec-1999

CLASSIFICATION: UNCLASSIFIED

CLASSIFICATION: ~~CONFIDENTIAL~~
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110 955

APPLICATION NUMBER: 89/1
FILING DATE: 04-DEC-1998

FILING DATE: 04-DEC-13
ATTORNEY/AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION:
NAME: CONN. NICHOLSON

NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27286

REGISTRATION NUMBER: 27386
 REFERENCE/DOCKET NUMBER: 960306 9E017

REFERENCE/DOCKET NUMBER: 5

TELECOMMUNICATION INFORMATION:
TELEPHONE (608) 251-5000

TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTER.

LENGTH: 31960

TYPE: nucleic acid

STRANDEDNESS: do

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCES

Alignment Scores:

Pred. No.:	4.08e-11	Length:	31960
Score:	188.50	Matches:	112
Percent Similarity:	33.47%	Conservative:	55
Best Local Similarity:	22.44%	Mismatches:	163
Query Match:	6.63%	Indels:	169
DB:	4	Gaps:	22

US-10-029-020-14 COPY 1760 2300 (1-541) X US-09-453-702B-11 (1-31960)

Qy 103 ArgIleLeuTyrAspGlnAlaGlyArg---ProSerLeuTrpSerPro----- 117

26629 CCGTACCGGTAATGATGATACCGGCCGGTGA CCGAGCTGGTCAACCCGAGGGGCTGGAC 26688

Qy 118 -----SerSerArgLeuAsnGlyValAsn 125
 | | | | |
 : : :

Db	26689	TACCGCTTTGAGTACGGGCAGGACCGTGTGACCATCAGGACAGCCTGAACCCGGCGGAG	26744
QY	126	ValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSerGlu---	144
Db	26749	GTGCTGTACACGGAAGCGGAG-----GGTGGCCTGAACACGTGTTGTGAAGAAGCAACAT	26802
QY	145	-----ArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIle	158
Db	26803	GCGGACGGGAGCATCACCCGACGAGTATGATGAGCGGGGAGGCTGAAGGCACACAGACG	26862
QY	159	PheAla-----	160
Db	26863	GATGCGCGGGACGGCGGACGGAGTACAGCCTGCATATGGCGTCGGGTGCGGTGACAGCG	26922
QY	161	-----AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeu	176
Db	26923	GTGACGGGGCGGACGGCAGGACGGTGGGTATGGCTAT-----	26961
QY	177	LeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSer	196
Db	26962	-----AACAGCCAGCGGCAG-----GTGACGCTCA	26985
QY	197	ValThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyr	216
Db	26986	GTGACGTACCCGACGGCTGCGCAGCAGC-----	27015
QY	217	ArgAsnIleTyrGlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAsp	236
Db	27016	-----CGGAGTATGATGAGAAAG	27033
QY	237	GlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGly	256
Db	27034	GGAAGGCTG-----	27042
QY	257	LysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAsp	276
Db	27043	-----GCGGCGGAGACCTCGCGCAGCGGAGAGACGACCGCGGTACAGTATGAT	27090
QY	277	GluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIle	296
Db	27091	GACCCGCGGAGTGAGCTGCCGACAGGGATACAGGACGCGACGGGCGAGTACAAAAACAGATG	27150
QY	297	ArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGly	316
Db	27151	GCATGGAGCCGTACGGT-----CAGCTGCTGACCTTTACGGACTGTCTCG	27195
QY	317	MetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGln	336
Db	27196	GGGTACACGACGCGGTATGATATGAC-----CGGTACGGTCAACAAATC	27240
QY	337	AlaValIleAsnGluThrProLeuProIle-----AspLeu	348
Db	27241	GCCGTTTACCCGGGAAGGATCAGCACTTACAGCAGTTATAACCCGCGTGGCCCACTG	27300
QY	349	TyrArgTyrAspAspValSerGlyLysThrGlu-----GlnPheGlyLysPheGlyVal	366
Db	27301	GTCAGTCAAGAGTGGCAGGGCGGTGAACCCGTTATGAGTACAGCGCCGCGCAGGC---	27357
QY	367	IleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHisPhe	386
Db	27358	-----GACCTCACCGCTATCGTTGCCCGGACGGCAGCCGAGTGAATACAGTAT	27408
QY	387	AspAlaTyrGlyArg--MetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr	405
Db	27409	GATGCGTGGGAAAGCCCGTCAGCACCCAGCGGGCGGTCTGACGCGCAGCATG-----	27462
QY	406	TrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysValGly	425
Db	27463	-----GGGTATGACGCTGCCGGGCGCATCACC-----GTGCTG	27495
QY	426	ProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThr	445
Db	27496	ACCAACGAGAACGGACGACGATCCACGTTCCGGTATGACCCCGGTGGACAGGCTGACTGAA	27555

```
QY 446 ValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHis 465
Db 27556 CAGCGGGTTTTGATGGCCGACGCAACGTTACCACTATGACCTGACCGGAAAA----- 27609
QY 466 LeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArg 485
Db 27610 CTCACGCAGAGTGAAGACGAGGGGCTGTCAACCTCTGGCACTACGATGCGTCGGACCGC 27669
QY 486 IleThrArgLeu-----GlyAsp-----ValGlnTyrLysMetAspGluAspGly 500
Db 27670 ATCACGCACCGACGGTGAACGGGACCCCGCAGACGACGTGGCAGTATGATGAGCACGGG 27729
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 27730 TGGCTA-----ACCACCCTCAGCCATACCACTGAGGCAAGGCCACCGGGTG--- 27771
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 27772 -----TCGGTCCACTACGGCTATGACGATAAAGGCCGC 27804
```

RESULT 9
US-09-328-352-1419
; Sequence 1419, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1419
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1419

Alignment Scores:
Pred. No.: 1.27e-12 Length: 2415
Score: 185.50 Matches: 142
Percent Similarity: 33.79% Conservative: 78
Best Local Similarity: 21.81% Mismatches: 232
Query Match: 6.52% Indels: 199
DB: 4 Gaps: 33

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-328-352-1419 (1-2415)

```
QY 33 AsnProThrValGlyLysArgAsnValThrLeuProIleAsp-----AsnGly 48
Db 220 AATTATGATTAGGTGATCAACATGTGGGTTTACCAATTGAAGTTGATGTGGAAATGGT 279
QY 49 LeuAsnLeuValGluTrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPhe 68
Db 280 AAAGAAATTACTCAATTGCTGGTTACAAAAGGGGTATTCCAACTCAAGTTAAACTTGCA 339
QY 69 GlyArgArgLeuArgValHisAsnArgAsnLeuLeuSerLeuAspPheAspArgValThr 88
Db 340 AATGGGGCGACAGAAACA-----AATATTGTAGAT---GACTTTGGAATATAACT 387
QY 89 ArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLeuTyrAspGln 108
Db 388 CAA-----CATACTGATGCAGATGGGGTTATTTCAAGAAAACAATATGATGAT 435
QY 109 AlaGlyArgProSerLeuTrpSerProSerArgLeuAsnGlyValAsnValThrTyr 128
Db 436 GCTGGCCGATTATATATAGATACACCAATTGTGGGTTAAATTACTCTACATTCACTTAT 495
QY 129 -----SerProGlyGlyTyrIleAlaGlyIleGlnArg 139
Db 496 GATGGTTTAACTGTTTCTAGGGTTGTAACTGGTGGCGGCAACTGAGCCGAATAGAAAA 555
```

```
QY 140 Gly-----IleMetSerGlu----- 144
Db 556 TATAATGGTGATGGTTTACTGATTTTCGAGTGAGGATAAGATATCAAATAAATCTATTATT 615
QY 145 ---ArgMetGluTyrAspGlnAlaGlyArgIleThrSerArg----- 157
Db 616 AACTCTAATAAGTATGATGATTTGGAAACCTGATATTTAAATCCAATCCTGGTTTTAGT 675
QY 158 IlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeu 177
Db 676 GCTATAACAAGTGGAACTACAAGCAGTTATGATGCTTTGACCGCCCAATTACTGTC--- 732
QY 178 LeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerVal 197
Db 733 -----AATGATAATGGAAGTGTGTG 753
QY 198 Thr-----MetProAsnValAlaArgGlnThr-----LeuGluThrIleArgSer 212
Db 754 ACATATTGCTATCAGTCTCTGTGGAGGAAAAAAGCTGGTGCAATTGTACAAAACAACAGATAGC 813
QY 213 ValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAsp 232
Db 814 TTTGGT-----ACAACAGAATCTAATCTATTAGCAGCAGGAGAT 852
QY 233 PheThrGluAspGlyHisLeuLeuHisThrPheTyrLeu-----GlyThrGlyArgArg 250
Db 853 TTTAGTGTGTGAT-----TTAAAAACACTGGTTGCACGTAAAGGTACAGATGGATCA 903
QY 251 ValIle----- 252
Db 904 GTCTTTTCAGACAACAACACTGAATTTGAAAACGCACTTTTAAAAACCAAAAGTTGCTGTATCG 963
QY 253 -----TyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyr 266
Db 964 GGTAGTTCTACACAAAGTTATACATATAACTTAATACAACTTTAGCA---ACAGAAAAA 1020
QY 267 AspThrThr-----LysValSerPheThrTyrAspGluThrAlaGlyMetLeuLys 283
Db 1021 GATAATAGTATTAGTGGGCAAAAGACATTTAAATATGATGATACT---GGCCGAATTACT 1077
QY 284 ThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyPro 303
Db 1078 TCAATTACCATCCAGATAGTTTCAGTAGAA---ACGATCAAAATATTTTCAATTAAGGAC 1134
QY 304 LeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAlaArgPheAsp 323
Db 1135 CTCATCGGTCGCGACATGGCGAGAGTTGAGACA-----ACT 1173
QY 324 TyrAsnTyrAspAsnSerPheArgVal---ThrSerMetGlnAlaValIleAsnGluThr 342
Db 1174 TATAGTTATTCTCTGCTGGCGGTTTAAAAACAACGACTAATGCATAATATTAGTGAAGCA 1233
QY 343 ProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGly 362
Db 1234 -----TTTGATCTTGATACATATGGCCGTTTATAAGTCATACACAAAAAATCAAT 1284
QY 363 -----LysPheGlyValIleTyrTyrAspIleAsn 372
Db 1285 GCCAATGATACCAATAATTCAATATGTCGTAAGATATGGC-----TATAATCAATTAAAC 1338
QY 373 GlnIleIleThr-----ThrAlaValMetThrHisThrLysHisPheAspAla 388
Db 1339 CAAGTTACTTCAATTCAATATCCTAATGGTAAATCTGTAATTTATCTAATCAGAATGCT 1398
QY 389 TyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr----- 405
Db 1399 TTAGGGAAGTAACATCAATACCA---AATGTAATACAAAGTTTGAATTATAATGCACGG 1455
QY 406 -----TrpMetThrValGln----- 410
Db 1456 CAACAACCTTACAACACTGTGCAAGCCCAATACCGATACTTTTATGGTCATATACTTATAATGAT 1515
QY 410 ----- 410
```


Db 1516 AGCGGGTTGCTGAATAATATCTCTGCTACAAGTCTGGAAAAGTGTCTCCTCAATGTTGAT 1575

QY 411 -----TyrAspAsnMetGlyArgValVallysLysGluLeuLysValGlyProTyrAla 428

Db 1576 TATGTTACGATAAAATTGAACCGAGTTAATAAGCTTTCGGACAAATGTGGATCAGTTTAT 1635

QY 429 AsnThrThr----- 431

Db 1636 AACGCTACAATTGACCGTTATGGTACAGGTTTGATGAGTACTGTTGAGTTAGATCAAGCA 1695

QY 432 ArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLys 451

Db 1696 CGTTATCAATATAGTTATAATAATGAT--GACATTACTAAAGTTAATATCATTCAAAA 1752

QY 452 -----ProLeuTrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHis 465

Db 1753 AGCTCGACAGTAGCACCAGCTATCTATACATATAACTAT--GTAATAACACTTCTCGA 1809

QY 466 LeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArg 485

Db 1810 TTAGCAAGCGTAAGTGGCTCAACT-----TATAGCTTTACTTATGATGCGATGGGTAAT 1863

QY 486 IleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArg 505

Db 1864 GTAACAAATGATGTTGCGTACTCTAACTTATGATAATTATAGTCGCGCTCAATAAGAAT 1923

QY 506 GlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAla 525

Db 1924 GGTAATGAGACCTATCTTTATAATGCTGATGGACTAAGAGTCCGTCGTGTAAGGGATGAT 1983

QY 526 GlySerTrpSerValArgTyrArgTyrAspGly 536

Db 1984 GGTTTAACAGATTATATTATGATCTTGATGGT 2016

RESULT 10

US-09-453-702B-90/c

; Sequence 90, Application US/09453702B

; Patent No. 6365723

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Valerie

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/453,702B

; FILING DATE: 03-Dec-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95017

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 90:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4453

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 90:

US-09-453-702B-90

Alignment Scores:

Pred. No.: 3.55e-12 Length: 4453

Score: 185.50 Matches: 116

Percent Similarity: 34.01% Conservative: 68

Best Local Similarity: 21.44% Mismatches: 220

Query Match: 6.52% Indels: 137

DB: 4 Gaps: 21

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-453-702B-90 (1-4453)

QY 49 LeuAsnLeuValGluTrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPhe 68

Db 1467 CTGGTGTGACCAACGAGCGCGAGCGGGCGGAAGAGGCCCGTAAACAG----- 1420

QY 69 GlyArgArgLeuArgValHisAsnArgAsnLeuSerLeuAspPhe----- 84

Db 1419 -----CACACCGCTTCTTTATCTTCCCTGACACACCCCGCCCTCTT 1378

QY 85 -----AspArgValThr 88

Db 1377 TCAGACTCAGCGTTCCCGACACACTGCCCGGTACCGAATACGGTCCCGACAGAGGTATC 1318

QY 89 ArgThrGluLysIleTyrAspHisArgLysPhe----- 100

Db 1317 CGCCTTTCCGGGTGTGGCTGACGCACGACCCCGCATACCCGAAAGCCTGCCCGTGGCG 1258

QY 101 ---ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSer 119

Db 1257 CCACTGGCGGTACACGTATACGGAAGCCGGTGAACCTGCTGGCGGTATATGACCGCAGC 1198

QY 120 ArgLeuAsnGlyValAsnValThrTyrSer-----ProGlyGlyTyrIleAlaGly 136

Db 1197 AATACGCAGGTGCGCGCTTTCACGTATAACGCGCAGCATCCGGGCCGATGGTGGCGCAC 1138

QY 137 IleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSer 156

Db 1137 CGTTACGCGGGAAGCCGGAGATGCGCTACCGCTACGACGATACCGGGCGGTGGAG 1078

QY 157 ArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValleu 176

Db 1077 CAGCTGAACCCCGCAGGCTGAGTTACCGCTACCGCTATGAGCAGGACCGCATCACCGTC 1018

QY 177 LeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSer 196

Db 1017 ACGGACAGCCTGAACCGCGGTGAGGTGCTGCATACAGAAGCGGGCGGCTGAAGCGG 958

QY 197 ValThrMetProAsnValAla----- 203

Db 957 GTGGTGAAAAAAGAACTGGCGACGGCAGCGTCAACGACAGCGGGTATGACGCGGCAGGA 898

QY 204 ArgGlnThrLeuGluThrIle-----ArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 897 AGGCTCACGGCGCAGACGACGCGGGGAGACGGCGGACAGAGTACCGTCTGATGTGGTG 838

QY 221 Gln-----ProProGluGlyAsnAlaSerValIleGlnAspPhe 233

Db 837 TCCGGCGATATACCGACATCACACCGGACGCGGGGAGACG-----AAATTTAC 784

QY 234 ThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGly-----ArgArg 250

Db 783 TATAACGACGGGAACCAAGTACGGCGGTGGTGTCCCGGACGGGCTGGAGAGCCGCCCG 724

QY 251 ValIleTyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLys 270

Db 723 GCATATGATGAACCGGCGAGCTG-----GTATCGGAGACATCGCGCTGTGGGACGTC 670
QY 271 ValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGlu 290
Db 669 ATCCGGTATGCTTATGATAATCCGACAGTGAATTACCGGCGACGACAACAGATGCGAGC 610
QY 291 GlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePhe 310
Db 609 GGCAGACCCGGCAGATGACGTGGAGCCGCTACGGG-----CAGTTGCTG 565
QY 311 ArgPheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsn----- 325
Db 564 GCGTTCACTGACTGCTCGGGCTACCAGACCCGTTATGATACGACCCGCTTCGGGCCAGATG 505
QY 326 -----TyrAspAsnSerPheArg 331
Db 504 ACGGCGGTTACCCGCGGAGGAGGTATCAGCCGTTACCGCGCTATGACAACCGTGGCCGG 445
QY 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351
Db 444 TTAACCTCGTGAAA----- 430
QY 352 AspAspValSerGlyLysThrGlu-----GlnPheGlyLysPheGlyValIleTyrTyr 369
Db 429 ---GACGCACAGGGCGGTGAACCGCGGTATGATGATAACACCGCCGACGGC----- 385
QY 370 AspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyr 389
Db 384 GACCTGACTGCCGTTATCACCCTCGGACGGCAACCGGAGCGAGACACAGTACGATCGGTGG 325
QY 390 GlyArg---MetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThr 408
Db 324 GGAAGAGCGGTGACGACCAACACAGGGCGGGCTGACCGCGAGTATG----- 280
QY 409 ValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAla 428
Db 279 ---GAGTATGACGCGCAGGACGTGTCCACCACGTGACCAACAG----- 238
QY 429 AsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIle 448
Db 237 AACGGCAGCCGAGTGAATTTACCTACGATGTGCTGGACCGCTGACGAGCAGCGCGC 178
QY 449 AsnAspLysProLeuTrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuSer 468
Db 177 TTTGATGGCAGCACACACGCTTACCGGTACAGCGCCACAGGG-----CAGCTCATCCGC 124
QY 469 ProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArg 488
Db 123 AGCAGGATGAGGGCAGGTACCGCAGTGGTACTACGATGAAGCAGACCGCATCACGCAC 64
QY 489 Leu-----GlyAsp-----ValGlnTyrLysMetAspGluAspGlyPheLeuArg 503
Db 63 CGGACGGTGAACGGCGACCCGCGAGCAGTGGCAGTATGATGAGCAGGGTGGTTACAG 4
QY 504 Gln 504
Db 3 AAA 1

RESULT 11

US-09-252-991A-11853
; Sequence 11853, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11853
; LENGTH: 2997
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11853
Alignment Scores:
Pred. No.: 2,82e-12 Length: 2997
Score: 184.00 Matches: 130
Percent Similarity: 37.14% Conservative: 88
Best Local Similarity: 22.15% Mismatches: 203
Query Match: 6.47% Indels: 167
DB: 4 Gaps: 29
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-11853 (1-2997)
QY 56 GlnArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArgArg----- 71
Db 531 CAACGGCGTCAACCAG-----CAGCCTCGCCTACACCGCGCTCGACGGCTGCTGGC 581
QY 72 LeuArgValHisAsnArg-AsnLeuLeuSerLeuAspPheAsp----- 85
Db 582 CTCGTGTCAGCACCGCGCGGCGAGCACCCCGCTTCGACTATGACGAGTCGCGCCAGATCAC 641
QY 86 -ArgValThrArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLe 105
Db 642 CCGCGTCAACCGGTGGCGACGGCAGTTGG-----CTGAGCTACGA 680
QY 105 uTyrAspGlnAlaGlyArgProSerLeu-Arg-----SerProSerSerArgLeuAsnGlyV 124
Db 681 ATACGACGACGCGCGCGC-----CTGGTGGCGATCGCAACCACTCGCGCGAACGCCT 734
QY 124 alAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArg----- 139
Db 735 CGAATACGACGCTGACACCAAGGCA---ATCGACC-GCCAGCGCATCAAGGACGCCA 790
QY 140 -----GlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerA 157
Db 791 GCGGCAGCCTGTGTGCGCCAGCAGCAATGGCGCTACGACGAGCTTGGCGCGCTGCTCGTG 850
QY 157 rgIlePheAlaAspGlyLysThrTrpSerTyrThrTyr-----L 170
Db 851 CGGTGCGCGCGCGCGCGCAGACACGACGAGCTTCGCTACGACCTCAACGACAATCCGGTCG 910
QY 170 euGluLysSerMetValLeuLeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspL 190
Db 911 GCGAAACCAACCCGCGCCAGTTGCGCCACAGCCAG-----GCCTTCGACG 955
QY 190 ysAsnAspArgLeuSerSerValThrMetProAsnValAlaArgGlnThrLeuGluThri 210
Db 956 CCCTCGACCGGTGGTGGCCAGAGCGATCCC----- 987
QY 210 leArgSerValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValI 230
Db 988 -----CTCGGGCGCAAGACCCGACTCGCCTACGACGCCAGGACAACCTCACCGAGG 1039
QY 230 leGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgA 250
Db 1040 TCAAGGAC-----CCGCGCGCGTCA 1060
QY 250 rgValIleTyrLysTyr-----GlyLysLeuSerLysLeuAlaGluThrLeuTyrA 267
Db 1061 CCACCGCTACGAATACGACGCGCTCGGCAACCTGACCCGACTGCTCAGCCCG-----G 1114
QY 267 spThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnL 287
Db 1115 ACAGCGGCACCAACCACTTCGAGCAGCAGCGCGCGGCAACGTCATCCGCGCACCGACG 1174
QY 287 euGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspA 307
Db 1175 CTCGC-----GGTGGGTGTCACCGAGTATCGCTACGACGCCCTCAATCGGCTGGTCGAGC 1228

```
QY 307 rgGlnIlePheArgPheThrGluGluGlyMet---ValAsnAlaArgPheAspTyrAsnT 326
|||
Db 1229 GCCGC-----TCGCCGAGCGACCCGAGCCTCGACGTACAGTACCGCTACGACCTCACCG 1282

QY 326 yrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluThrProLeuProI 346
|||
Db 1283 CCGACGGCAACACGAGGCATCGCGCGCTGGCGCCATC-----GAAGGTGCCCGCGACA 1336

QY 346 leAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGlnPheGlyLysPheGlyV 366
|||
Db 1337 GCCTGGTGTACCGCTACGACGAGCGCGGCAACCTGGTTCGAGCAGGTACGCAGC----- 1389

QY 366 alileTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHisP 386
|||
Db 1390 --ATCCGCCTCGACACGACGACCTGTCTCGACCGCGTG-----ACCTACCGCT 1435

QY 386 heAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr- 405
|||
Db 1436 ACGACGCGGCCAACCAACTGTCTGGAGATCGGCTACCCCTCCGGCCTCGCCATCGGCTACC 1495

QY 405 ----- 405

Db 1496 CGCGCAACGCCGCGGCCAGGTCCGCCAGCTGCAGTGGCGGCAAGCGGCCGA 1555

QY 406 -----Trp- 406
|||

Db 1556 GCACCCCTGGTTGGGCAGATCGCCTACCTGCCCTTCGGCCCGCTGCAGCGCCTGACCTGGG 1615

QY 407 -----MetThrValGlnTyrAspAsnMetGlyArgValVallyLysGluL 422
|||
Db 1616 GCAACGGCATCACTCTCAGCCGCGAGTACGACCGAGACTACCAACTGTCTGGCGCAG---- 1671

QY 422 euLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyG 442
|||
Db 1672 --AAGTCCGGCCCTTGGCAGAGC-----GACTACCAGCAGCATGCCAATGGCA 1717

QY 442 lnLeuGln-----ThrValSerIleAsnAspLysProLeuT 454
|||
Db 1718 ATATCCAGCAGCACCGCCACAGCCTCTGCGGCAACCTTGGACTACCAGTACGACCCGCTGG 1777

QY 454 rpArg-----TyrSerTyrAspLeuAsnG 462
|||
Db 1778 ACCGCTGACCGAGGAACGCGGCGTCCAGGGCGGCGCAGCTACGCCTACGACGCGGTCCG 1837

QY 462 lyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspI 482
|||
Db 1838 GCAACCGCACCCCAACGCAGCAGCAACCCCGCTCAGGCGGCAACCGCAGCAGCCAGACT 1897

QY 482 leArg-----AspArgIleThrArgLeuGlyAspValGlnTyrLysMetA 497
|||
Db 1898 ACCAGTACGCGCGCGACAGCAACCGATTGACCGCCATCGCGCGCGCAAGCGGTGACCGAG 1957

QY 497 spGluAspGlyPheLeuArgGln--ArgGlyGlyAspIlePheGluTyrAsnSerAlaG 516
|||
Db 1958 ACCCGCGCGCAACCTCACCCAGGACCGCGCGCGCAAACTGGCCTACGACGCCCGAG 2017

QY 516 lyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTyrPheValArgTyrArgTyrAspG 536
|||
Db 2018 GCCGCTG---CAAAGCGTCAGCCTCGACGCGCCAGCAGGTGCGCCGAATACCGCTACACG 2074

QY 536 lyLeuGlyArgArgVal 541
|||
Db 2075 CCCTCGGCGCAGCGCATC 2091
```

RESULT 12

US-09-252-991A-7069
; Sequence 7069, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7069  
; LENGTH: 2925  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7069  
  
Alignment Scores:  
Pred. No.: 4.17e-12 Length: 2925  
Score: 182.50 Matches: 115  
Percent Similarity: 36.06% Conservative: 88  
Best Local Similarity: 20.43% Mismatches: 203  
Query Match: 6.41% Indels: 157  
Db: 4 Gaps: 26
```

```
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-7069 (1-2925)  
  
QY 106 TyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArgLeuAsnGlyVal--- 124  
|||  
Db 493 TATGACGAGCGCGCAAGCCCGGCGAGCATCACCGATGCC-----AACGGCGTCACC 543  
  
QY 125 ---AsnValThrTyrSer---ProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMet 142  
|||  
Db 544 AGCAGCCTCGCCTACACCGCGCTCGACGGCTGGTGGCTCCGTACGACACCGCGGCAGC 603  
  
QY 143 SerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGly 162  
|||  
Db 604 ACCACCGCTTCGACTATGACGGCTCGGCGCAGATCACCGCGTACCGCGCGGCGACGGC 663  
  
QY 163 LysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSerGlnArg 182  
|||  
Db 664 AGTTGGCTGAGCTACGAATACGACGCGCGCGCTGGTAGCGATCGGCAACAACCTC 723  
  
QY 183 GlnTyrIlePheGluPheAsp-----LysAsnAspArgLeu----- 194  
|||  
Db 724 GCGCAACGCTCGAATACGACGCTCGACACCAAGGGCAATCGCACCGCGCGCATCAAG 783  
  
QY 195 -----SerSerValThrMetProAsnValAlaArgGlnThrLeu---GluThr 209  
|||  
Db 784 GACGCCAGCGCGCCTGGTGGCGCCAGCAATGGGCCTACGACGAGCTTGGCCCGGTG 843  
  
QY 210 IleArgSerValGly-----TyrTyrArgAsnIleTyrGlnPro 222  
|||  
Db 844 CTCGCTGCGTTCGGCGCGCGCGCCAGACACGACGAGCTTCGCTACGACCTCAACGACAAT 903  
  
QY 223 ProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThr 242  
|||  
Db 904 CCGTTCGGCGAA-----ACCAACCGCGCGCAGTTCGCCCCACAGC 942  
  
QY 243 PheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLysLeuAla 262  
|||  
Db 943 CAGCGCTTCGATGCCCTCGACCGGTGGTGGCCAGAGCGATCCCTCGCGCGCAAGACC 1002  
  
QY 263 GluThrLeuTyrAsp-----Thr 268  
|||  
Db 1003 CGACTCGCCTACGACGCGCCAGACAACCTCACCGAGGTCAAGACCGCGCGGTGTCAAC 1062  
  
QY 269 ThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGln 288  
|||  
Db 1063 ACCCGC-----TACGAATACGACGGCCTCGGCAACCTGACCCGACTGGTCAGC----- 1110  
  
QY 289 AsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGln 308  
|||  
Db 1111 CCGGACAGCGGCACCAACCACTTCGAGCAGCAGCGCGCGCGCAACGTCATCCGCCGCCACC 1170  
  
QY 309 IlePheArgPheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsn 328
```


Db 1171 GACGCTCGC-----|||:::||||| 1209

QY 329 SerPheArgValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeu 348

Db 1210 CTCAAATCGGCTGGTCAGCGCGCTCGCGG--AGCGACCGGAGCTCGACGTACAG-- 1263

QY 349 TyrArgTyrAspValSerGlyLysThrGluGlnPheGlyLysPheGlyValIle--- 367

Db 1264 TACCGCTACGACCTCACCGCGACGCGCAACAAGGGCATCGCGCGCTGGCGCCATCGAT 1323

QY 368 -----TyrTyrAsp-----IleAsnGlnIle 374

Db 1324 GCGGCTCGGACAGCCTGGTGTTACCGCTACGAGCGCGGCAACCTGGTCGAGCAGGTG 1383

QY 375 IleThr-----ThrAlaValMetThrHisThrLysHisPheAspAla 388

Db 1384 CGCAGCATCGCCTCGACCGACAGACCCCTGCTCGACCGCGTGACCTACCGCTACGACCG 1443

QY 389 TyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr----- 405

Db 1444 GCCAACCAACTGCTGGAGATCGGCTACCCCTCCGGCCTCGGCATCGGCTACCCGCGCAAC 1503

QY 405 ----- 405

Db 1504 GCGGCGGCGCAGGTGCGCCAGCGTGACCCCTGGCAGTGGCGCAAGCGCGGACACCCCTG 1563

QY 406 -----Trp----- 406

Db 1564 GTCGGGCAGATCGCCTACCTGCCCTTCGGCCCGCTGCTGGCGCTGACCTGGGGCAACGCG 1623

QY 407 -----MetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysVal 424

Db 1624 ATCACTCTCAGCCGCGAGTACGACGAGGACTACCAACTGCTGCGGCAG-----AAGGTC 1677

QY 425 GlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGln 444

Db 1678 GGCCTCTGGCAGAGC-----GACTACGACGACGATGCCAATGGCAATATCCAG 1725

QY 445 -----ThrValSerIleAsnAspLysProLeuTyrArg--- 455

Db 1726 CAGCACCGCCACAGCCTCTGGGCGACCCCTGGACTACGACTACGACCGCGCTGGACCGCCTG 1785

QY 456 -----TyrSerTyrAspLeuAsnGlyAsnLeu 464

Db 1786 ACCGAGGAACGCGCGTCCAGGCGGCGCAGCTATGCCTACGACGCGGTGCGGCAACCGC 1845

QY 465 HisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArg--- 483

Db 1846 ACCCAACGACGCGACAACCCCGCTCAGGCGGCACCGCCGACGAGGACTACCAGTAC 1905

QY 484 -----AspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499

Db 1906 GCGCCGACAGCAACCGATTGACCGCCATCGCGCGCAAGCGGTGACCGACGCGCGCC 1965

QY 500 GlyPheLeuArgGln---ArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeu 518

Db 1966 GGCAACCTCACCGACCGCGCGCGCGCAAACTGGCCTACGACGCCCGCGCCCTG 2025

QY 519 IleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGly 538

Db 2026 ---CAAGAGCTCAGCCTCGACGCGCAGCAGGTGCGCCGAATACCGCTACAACGCGCTCGGC 2082

QY 539 ArgArgVal 541

Db 2083 CAGCGGATC 2091

RESULT 13

US-09-328-352-2478

; Sequence 2478, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 2478

; LENGTH: 4884

; TYPE: DNA

; ORGANISM: Acinetobacter baumannii

US-09-328-352-2478

Alignment Scores:

Pred. No.: 1.14e-11 Length: 4884

Score: 182.00 Matches: 137

Percent Similarity: 30.87% Conservative: 89

Best Local Similarity: 18.72% Mismatches: 198

Query Match: 6.40% Indels: 308

DB: 4 Gaps: 34

US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-328-352-2478 (1-4884)

QY 80 LeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspHisArgLys 99

Db 1822 ATGAACCTAAATGGGATGGTATTTTACCTACTTCAAAAGCCATTGAAGAAATGGGCC--- 1878

QY 100 PheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeu-----TrpSerPro 117

Db 1879 -----GACAATGCTAGTCTGCTTCAAAGCTGGAGTGGGATAAA 1917

QY 118 SerSerArgLeu-----AsnGlyValAsnValThrTyrSer 129

Db 1918 AATATCCGAAAAACGACAGCTGTAGACGTTGAGGGGAATAGCACTGAACATTACTACGAT 1977

QY 130 ProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGlu----- 147

Db 1978 ATTGATGGCTAT-----ACTTACCGTATTGTTTATCTGTATAATTTTGAAGAGTGT 2028

QY 148 -----TyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGlyLysThrTrp 165

Db 2029 TTCCTTCGAGATGATGCGAAAAATATAACACTCCATATTCCTAAAGACGGTTCGAAAACT 2088

QY 166 SerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSerGln-----ArgGln 183

Db 2089 TCTTACACTTATGATGAGCGTGGTAATGTCTTGACTACAACCTCAAGATGATGGTGGCACA 2148

QY 184 TyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetProAsnValAla 203

Db 2149 AGTTATTTGAATATGATGAAAAAATCAACTCACTGGTGTGTGATGCCGAACAGGGA 2208

QY 204 Arg----- 204

Db 2209 CGTGGTTTAAAGCAATATGATGGTTCTGGAAACCTGATCAAGAAATATGATCCACTTAAG 2268

QY 205 ---GlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProPro 223

Db 2269 CATGAAACTGCTTATGTTTACAATGCGATGGGCGCTAGTCACTTCCATTACAGATGCTAAA 2328

QY 224 GluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPhe 243

Db 2329 GGTGGAAGTAAGTCACTTAATATATGAT-----GATCAAGGGAATTTAATT---AGTTAT 2379

QY 244 TyrLeuGlyThrGlyArgArgValIleTyrLysTyr-----GlyLysLeuSerLys 260

Db 2380 ACAGATTGCTCAGGTAAAGAGACCAAGTGGCAATATGATGAGCGTGGTCTGTATAATAAGC 2439

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrThrTyrAspGluThrAla--- 279

Db 2440 ATTGAAAAATGCATTA-----AACCAAAAAAGTTGAGTATTTTATACTGAACCTTACCTTA 2493

QY 280 -----GlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThr 295

Db 2494 GAAAAATCGGAACCAATTATAAAGGGTTTACCTCTGAATGCTTTTGGTCAACTTGAAAAAG 2553

Qy 296 IleArgTyr----- 298
Db ATTAACATGCTGATGGCACAGAGAACACTTCATTATGATGCTGAAGGCGTTTACTG 2613
Qy 299 ArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetVal 318
Db 2614 GCGATGTAGATCCTAAACAGACATCACTCGCTATGAATATGATGAGCGAGGTTTAAT 2673
Qy 319 AsnAlaArgPheAsp-----TyrAsnTyrAspAsnSerPheArg 331
Db 2674 CTGTCTCGTACAGATGCTCTAAACCATAAATTAATAATATAAATGGGATCGATTGGGACGT 2733
Qy 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351
Db 2734 CTCACCGGACTC-----ATTAATGAAATGGTGCCAGTTACCAATTCTTC---TAT 2781
Qy 352 AspAspValSerGlyLysThrGluGlnPhe-----GlyLysPheGlyValIleTyr 368
Db 2782 GATGTCGCAAGCCGACTCGTTAAAGAAATTGACTTTGATGGTAAGGAAACGGTTTACCAT 2841
Qy 369 TyrAsp-----IleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHisPhe 386
Db 2842 TACGATGAAAGTCGGGCCAACTGCAACCACTATTGAAGTTGCA----- 2886
Qy 387 AspAlaTyrGlyArg---MetLysGlu---ValGlnTyrGlu 398
Db 2887 TCTGCTTATGGACAAGACCTCAAGACCGGAGCTGCACCGAAAGACCGAATTCACAAATTT 2946
Qy 399 IlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVal 418
Db 2947 ATTTTT-----GACAGTATGGGCGTTTAGAA 2973
Qy 419 LysLysGluLeuLysValGlyProTyrAla-----AsnThrThrArg 432
Db 2974 CAGCGTACAGCGGGTTATGGTCTATGTTAGAACTTGAAGAAAAACAGACTGAAGAA 3033
Qy 433 TyrSerTyrGlu----- 436
Db 3034 TTTGCTATGATTATATATGGGGCGAATCATTCAGCTAAATATGCTCAAAGTAATTTACAA 3093
Qy 437 ---TyrAspAlaAspGlyGlnLeu-----GlnThrValSerIleAsn 449
Db 3094 TGGTTTATGATGCGGCTGGCAATTAGTACAGGAACACCAGCAAGATTATAAATCAAT 3153
Qy 450 AspLysProLeuTrpArgTyrSerTyrAsp----- 459
Db 3154 AAACAGCGGTGTGGAACACCAATATGACGAAATTAATGATCGAATCAAAACAACTCGT 3213
Qy 459 ----- 459
Db 3214 CCAGACGGGAGGTTATTGACTGGTTGACCTATGGTAGCGGTGATGTACAAAGTTTAATA 3273
Qy 460 LeuAsnGly-----AsnLeuHis----- 465
Db 3274 GTGAATGGTCAAGACTTTGTAGCTTTGAGCGTGTGATGATTATACACCGTGAAATCGCACGT 3333
Qy 465 ----- 465
Db 3334 CACTATGCCAACGGAGTGAGTCAAGAACAGCAATATGACTTAGCAGGTGCTTTGAAAAGC 3393
Qy 466 -----LeuLeuSerPro-----GlyAsn 471
Db 3394 CAAATGATGTTAAGTGAACATGAAAATGGCTATCAAAACCAATATAAGCGCCACAACAAT 3453
Qy 472 SerAlaArgLeuThrPro-----LeuArgTyrAsp----- 481
Db 3454 GCATTAGAGCAGACCAGTCACTGGTACAGCGTCTTTTATCAATATGACAAACAGGTGAG 3513
Qy 482 -----IleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGlu 498
Db 3514 CTAACGGCAATTCGAGAT-----ACACGCCGTGGCAATATTGCTTATAAATACGATCCT 3567

Qy 499 AspGlyPheLeuArgGlnArg-----GlyGlyAspIlePheGluTyrAsnSer 514
Db 3568 GTAGGCGGTTTACTTGAAGCGAGCAGTAAGTTAGGTAAGAAACATTTAGTTTGGACCCT 3627
Qy 515 AlaGlyLeuLeuIleLysAlaTyr----- 522
Db 3628 GCGAGTAACATTCTAGATTCTATACCATAGTCAGAAAAGTTTCAAGCCATTTCACAAAAGCTA 3687
Qy 522 ----- 522
Db 3688 GACGAAACGAGTTATGGTTTATAATCGTTTGGTTAATAATAGTTAAAGATATTTAGAC 3747
Qy 522 ----- 522
Db 3748 CAACAATACCAATATGATGCTTATGGACAACACTCATACGACAAAAACGAGCCAGGTGAT 3807
Qy 523 -----AsnArgAlaGlySerTrpSer 529
Db 3808 TTAACCTTGAGTGGGATGTCTATGGCCGCATGGTAAAGAGCGGTAAATAGTCAATACACG 3867
Qy 530 ValArgTyrArgTyrAspGlyLeuGlyArgArgVal 541
Db 3868 GCAGAAATATCGTTATGATGCTTTAGGCCGCCGTATT 3903
RESULT 14
US-09-252-991A-11787
; Sequence 11787, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11787
; LENGTH: 1857
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11787
Alignment Scores:
Pred. No.: 8.26e-12 Length: 1857
Score: 177.50 Matches: 113
Percent Similarity: 35.44% Conservative: 89
Best Local Similarity: 19.82% Mismatches: 196
Query Match: 6.24% Indels: 172
Gaps: 24
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-11787 (1-1857)
Qy 54 TrpArgGlnArgLysGluGlnAlaArgGlyGlnValThrValPheGlyArgArgLeuArg 73
Db 31 TGGCGACGGCAGTTGGCTGAG-----CTACGA 57
Qy 74 ValHisAsnArgAsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIle 93
Db 58 ATACGACGACGACGCGCGCTGGTGGCGATCGGCAACAACTCGGCGAACCGCTCGAATA 117
Qy 94 TyrAspAspHisArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSer 113
Db 118 CGACGTCGACACCAA-----GGGCAATCGCACCGCCACG 153
Qy 114 -LeuTrpSerProSerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTy 133
Db 154 CATCAAGGACGCCAGCGCGCAGCTG----- 178
Qy 133 rIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyAr 153

Db	179	-----GTGGCCAGCAGCAATGGGCCTACGACGAGCTTGCCG	216
QY	153	gileThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyr	169
Db	217	GCTGCTCCGTGCGGTGCGCGCGCGGCAGACACGAGCTTCGCCTACGACCTCAACGA	276
QY	170	-----LeuGluLysSerMetValLeuLeuLeuHisSerGlnArgGlnTyrIlePh	186
Db	277	CAATCCGGTCGGCGAAACCAACCCGCGCCAGTTTCGCCACAGCCAG	322
QY	186	eGluPheAspLysAsnAspArgLeuSerValThrMetProAsnValAlaArgGlnTh	206
Db	323	-GCCTTCGACGCCCTCGACCGGCTGGTCGGCCAGAGCGATCCC	364
QY	206	rLeuGluThrIleArgSerValGlyTyrTyrArgAnIleTyrGlnProProGluGlyAs	226
Db	365	-----CTCGCGGCAAGACCCGACTGCGCTACGACGCCCAGGACAA	405
QY	226	nAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGl	246
Db	406	CCTCACCGAGGTCAAGGAC	426
QY	246	yThrGlyArgArgValIleTyrLysTyr	263
Db	427	GCGCGGCGTCAACCCGCTACGAATACGACGGCCTCGGCAACCTGACCCGACTGGTCAG	486
QY	263	uThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLy	283
Db	487	CCCG-----GACAGCGGCACCAACCACTTCGAGCAGCAGCCGCGGCAACTCATCCG	540
QY	283	sThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyPr	303
Db	541	CGCACCGACGCTCGC-----GGTGGGTACCGAGTATCGTACGACGCCCTCAATCG	594
QY	303	oLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMet	322
Db	595	GCTGGTCGAGCGCCG-----TCGCCGAGCGACCCGAGCCTCGAGTACGACCGCTA	648
QY	322	eAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluTh	342
Db	649	CGACCTCACCGCGCACGGCAACCAAGGCATCGGCGCCTGGCGCCATC-----GAAGG	702
QY	342	rProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGl	362
Db	703	TGCCCGGACAGCGTGTGTACCGCTACGACGAGCGCGGCAACCTGGTCGAGCAGGTACG	762
QY	362	yLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHi	382
Db	763	CAGC-----ATCCGCTCGACCAAGCAGACCCCTGCTCGACCGCGTG-----	802
QY	382	sThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSe	402
Db	803	-ACCTACCGCTACGACGCGGCAACCAACTGTGGAGATCGGCTACCCCTCCGCGCTCGC	861
QY	402	rLeuMetTyr	405
Db	862	CATCGGCTACCCGGCAACGCGCGGCCAGGTGCGCCAGGTGACCCCTGGCAGTGGCGGA	921
QY	405	-----	405
Db	922	CAAGGCGCGAGCACCCCTGGTTGGGAGATCGCCTACCTGCCCTTCGGCCCGCTGCAGCG	981
QY	406	-----Trp-----MetThrValGlnTyrAspAsnMetGlyArgValVa	418
Db	982	CCTGACCTGGGGCAACGGCATCACTCTCAGCGCGAGTACGACCAAGGACTACCAACTGCT	1041
QY	418	lLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAs	438
Db	1042	GCGGCAG-----AAGTCGGCCCTGGCAGAC-----GACTACCAAGCACGA	1083
QY	438	pAlaAspGlyGlnLeuGln-----ThrValSerIleAsnAs	450

Db	1084	TGCCAATGGCAATATCCAGCAGCACCAGCCACAGCCCTCTGGGGCACCCCTGGACTACCACTA	1143
QY	450	plysProLeuTrpArg	458
Db	1144	CGACCCGCTGGACCGCCTGACCGAGGAACGCGGCTCCAGGGCGGCGCAGCTACGCCTA	1203
QY	458	rAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProle	478
Db	1204	CGACGCGTTCGGCAACCGCACCCCAACGCGACACCCCGCCTCAGCGGCGCACCGCCAG	1263
QY	478	uArgTyrAspIleArg-----AspArgIleThrArgLeuGlyAspValGl	493
Db	1264	CAGCCAGGACTACCACTACGCGCGCCGACAGCAACCGATTGACCGCCATCGGCGCGCAAGC	1323
QY	493	nTyrLysMetAspGluAspGlyPheLeuArgGln--ArgGlyGlyAspIlePheGluTy	512
Db	1324	GGTGACCAGCGACGCGCGCGCAACCTCACCCAGACCGCGCGCGCAAACTGGCCTA	1383
QY	512	rAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTy	532
Db	1384	CGACGCCAGGCGCGCCTG---CAAAGCTCAGCCTCGACGGCCAGCAGGTGCGCCGAATA	1440
QY	532	rArgTyrAspGlyLeuGlyArgArgVal	541
Db	1441	CCGCTACAAACGCCCTCGGCCAGCGCATC	1468
RESULT 15			
US-09-252-991A-7125			
; Sequence 7125, Application US/09252991A			
; Patent No. 6551795			
; GENERAL INFORMATION:			
; APPLICANT: Marc J. Rubenfield et al.			
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS			
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
; FILE REFERENCE: 107196.136			
; CURRENT APPLICATION NUMBER: US/09/252,991A			
; CURRENT FILING DATE: 1999-02-18			
; PRIOR APPLICATION NUMBER: US 60/074,788			
; PRIOR FILING DATE: 1998-02-18			
; PRIOR APPLICATION NUMBER: US 60/094,190			
; PRIOR FILING DATE: 1998-07-27			
; NUMBER OF SEQ ID NOS: 33142			
; SEQ ID NO 7125			
; LENGTH: 1857			
; TYPE: DNA			
; ORGANISM: Pseudomonas aeruginosa			
US-09-252-991A-7125			
Alignment Scores:			
Pred. No.:	1.27e-11	Length:	1857
Score:	176.00	Matches:	112
Percent Similarity:	35.30%	Conservative:	85
Best Local Similarity:	20.07%	Mismatches:	188
Query Match:	6.19%	Indels:	173
DB:	4	Gaps:	24
US-10-029-020-14_COPY_1760_2300 (1-541) x US-09-252-991A-7125 (1-1857)			
QY	72	LeuArgValHisAsnArgAsnLeuSerLeuAspPheAspArgValThrArgThrGlu	91
Db	52	CTACGAATACGACGACGACGCGCGCTGGTAGCATCGGCAACAACTCGGCGAAGCGCT	111
QY	92	LysIleTyrAspAspHisArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArg	111
Db	112	CGAATACGACGTCGACACCAA-----GGGCAATCGCACCGC	147
QY	112	ProSer-LeuTrpSerProSerSerArgLeuAsnGlyValAsnValThrTyrSerProGl	131
Db	148	CCAGCGCATCAAGGACGCGGCGCAGCTG-----	178
QY	131	yGlyTyrIleAlaGlyIleGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAl	151
Db	179	-----GTGCGCCAGCAGCAATGGGCCTACGACGAGCT	210

QY 151 aGlyArgIleThrSerArgIlePheAlaAspGlyLysThrTrpSerTyrThrTyr----- 169
Db 211 TGGCCGGCTGCTCGTTCGGTTCGGCGCGCGCCAGACACGACGCTTCGCCTACGACCT 270
QY 170 -----LeuGluLysSerMetValLeuLeuLeuHisSerGlnArgGlnTy 184
Db 271 CAACGACAATCCGGTCGGCGAACAACCCGCCAGTTCCGCCACAGCCAG----- 322
QY 184 rIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetProAsnValAlaAr 204
Db 323 -----GCCTTCGATGCCCTCGACCCGGCTGCTCGGCCAGAGGATCCC----- 364
QY 204 gGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProGln 224
Db 365 -----CTCGGGCGCAAGACCCGACTCGCCCTACGACGCCCA 399
QY 224 uGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTy 244
Db 400 GGACAACCTCACCGAGGTCAAGGAC----- 424
QY 244 rLeuGlyThrGlyArgArgValIleTyrLysTyr-----GlyLysLeuSerLysLe 261
Db 425 -----CCGCGCGGTGTACACACCCGCTACGAATACGACGGCCCTCGGCAACCTGACCCGACT 480
QY 261 uAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGlyMe 281
Db 481 GGTACGCCCG-----GACAGCGGCACCAACCACTTCGAGCACGACGCGCGCGGCAACGCT 534
QY 281 tLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnI 301
Db 535 CATCCGCGCGACCGACGCTCGC-----GGTGGGTTCACCGAGTAGTATCGCTACGACGCCCT 588
QY 301 eGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMet---ValAsnAl 320
Db 589 CAATCGGCTGGTCGAGCGCGC-----TCGCCGAGCGACCCGAGCCTCGACGTACAGTA 642
QY 320 aArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAs 340
Db 643 CCGCTACGACCTCACCGCGGACGGCAACAAGGGCATCGGCGCGCTGGGCGCC----- 694
QY 340 nGluThrProLeuProIleAsp-----LeuTyrArgTyrAspAspVa 354
Db 695 -----ATCGATGGCGCTCGCGACAGCGTGGTGTACCGCTACGACGAGCG 738
QY 354 lSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnI 374
Db 739' CGGCAACCTGGTCGAGCAGGTGCGCAGC-----ATCCGCCCTCGACCGACGACCCCT 789
QY 374 eIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGl 394
Db 790 GCTCGACCGCGTG-----ACCTACCGCTACGACGCGGCGCAACCAACTGCTGGA 837
QY 394 uValGlnTyrGluIlePheArgSerLeuMetTyr----- 405
Db 838 GATCGGCTACCCCTCCGGCGCTCGCCATCGGCTACCCGGCGCAACGCGGCGCGGCTCGC 897
QY 405 ----- 405
Db 898 CAGCGTGACCCTGGCAGTGGGCGACAAGGCGCCGAGCACCCCTGGTCGGGCGAGATCGCCTA 957
QY 406 -----Trp-----MetThrValGl 410
Db 958 CCTGCCCTTCGGCCCGCTGCTGGCCCTGACCTGGGGCAACGGCATCACTCTCAGCCCGCA 1017
QY 410 nTyrAspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnTh 430
Db 1018 GTACGACACGAGGACTACCAACTGCTGCGGCAG-----AAGTTCGGCCCTTGGCAGAGC-- 1069
QY 430 rThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGln----- 444
Db 1070 -----GACTACGACGACGATGCCAATGGCAATATCCAGCAGCACCGCCACAGCCT 1119

QY 445 -----ThrValSerIleAsnAspLysProLeuTrpArg----- 455
Db 1120 CTGGGGCACCTGGACTACAGTACGACCCGCTGGACCGCCTGACCGAGGAACGCGCGT 1179
QY 456 -----TyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGl 470
Db 1180 CCAGGGCGGCGCAGCTATGCCTACGACCGGGTTCGGCAACCCGACCCCAACGACGACGACAA 1239
QY 470 yAsnSerAlaArgLeuThrProLeuArgTyrAspIleArg-----AspAr 485
Db 1240 CCCCCTCAGCGGCACCCGAGCAGCCAGGACTACAGTACGCGCGCCGACAGCAACCG 1299
QY 485 gIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGln-- 504
Db 1300 ATTGACCGCCATCGGCGCGCAAGCGGTGACGACGCGCGCGCAACCTCACCAGGA 1359
QY 505 -ArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnAr 524
Db 1360 CCGCGCCCGCGCAAACTGGCCTACGACGCGCCAGGCGCGCTG---CAAAGCGTCAGCCT 1416
QY 524 gAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgVal 541
Db 1417 CGACGGCCAGCAGGTGCGCCGAATACCGCTACAAACGCCCTCGGCGCAGCGCATC 1468

Search completed: August 14, 2004, 21:39:25
Job time : 275.054 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 6344.74 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
Perfect score: 2845
Sequence: 1 YYIGADGSLRLLANGMEVA.....YNRAGSWSVRYRDGLRRV 541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020_@CGN_1_13135_@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1.	2834	99.6	6246	29	AY413475	AY413475 Homo sapi
2	2629	92.4	5970	29	AY413476	AY413476 Pan trogl
3	2233	78.5	5094	29	AY405420	AY405420 Homo sapi
4	2228	78.3	5069	29	AY405421	AY405421 Pan trogl
5	2228	78.3	5087	29	AY405422	AY405422 Mus muscu
6	2017.5	70.9	3038	11	AK037897	AK037897 Mus muscu
7	1177	41.4	894	13	BU151768	BU151768 AGENCOURT
8	1174	41.3	834	13	BU610605	BU610605 UI-M-FC0-
9	1168.5	41.1	951	13	BU839812	BU839812 AGENCOURT
10	1142	40.1	694	14	CF723353	CF723353 UI-M-GV0-
11	1112.5	39.1	758	14	CB723840	CB723840 UI-M-FY0-
12	1103	38.8	757	10	BE741784	BE741784 601595546
13	1099	38.6	868	29	CNS03W3H	AL263222 Tetraodon
14	1097	38.6	856	14	CF534506	CF534506 UI-M-GI0-
15	1068	37.5	1031	29	CNS02PHB	AL207992 Tetraodon
16	1011	35.5	675	14	CF540583	CF540583 UI-M-GV0-
17	1005	35.3	805	13	BU215426	BU215426 603756042
18	1000.5	35.2	868	14	CA454304	CA454304 AGENCOURT
19	999	35.1	812	9	AU133387	AU133387 AU133387
20	992.5	34.9	728	14	CF735545	CF735545 UI-M-HB0-
21	986	34.7	600	14	CA527167	CA527167 8032-72 M
22	983	34.6	690	28	AZ329056	AZ329056 1M0053E09
23	979	34.4	811	9	AU119933	AU119933 AU119933
24	961	33.8	655	14	CD351410	CD351410 UI-M-GI0-
25	954	33.5	1048	29	CNS03Z5U	AL267195 Tetraodon
26	938	33.0	819	13	BQ443891	BQ443891 UI-M-EW0-
27	920	32.3	944	29	CNS02T8X	AL212874 Tetraodon
28	896	31.5	775	13	BU611301	BU611301 UI-M-FI0-
29	888	31.2	751	14	CF535325	CF535325 UI-M-GH0-
30	885	31.1	764	13	BU328309	BU328309 603495004
31	876	30.8	592	10	BF349559	BF349559 MR0-HT015
32	874	30.7	668	13	BU214946	BU214946 603106244
33	860	30.2	763	29	CNS03WB1	AL263494 Tetraodon
34	844	29.7	589	14	CB434224	CB434224 610820 MA
35	837	29.4	882	13	BU914563	BU914563 AGENCOURT
36	831.5	29.2	900	13	BU267121	BU267121 603506284
37	826	29.0	675	10	BE374675	BE374675 601224176
38	817	28.7	1103	13	BX457474	BX457474 BX457474
39	808	28.4	633	14	CF728295	CF728295 UI-M-HB0-
40	799	28.1	577	29	DR7F15T	AL740238 Danio rer
41	778	27.3	748	14	CF723503	CF723503 UI-M-GV0-
42	765	26.9	444	14	CB546578	CB546578 AMGNNUC:N
43	764	26.9	843	9	AU124680	AU124680 AU124680
44	763.5	26.8	592	10	BE161779	BE161779 MR3-HT044
45	763	26.8	974	13	BQ425244	BQ425244 AGENCOURT

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AY413475
Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY413475
AY413475.1 GI:39769437
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6246)

AY413475
6246 bp
DNA
linear
GSS 12-DEC-2003

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 6246)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source Location/Qualifiers
1..6246
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>6246
/locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 6246
Score: 2834.00 Matches: 541
Percent Similarity: 99.63% Conservative: 0
Best Local Similarity: 99.63% Mismatches: 0
Query Match: 99.61% Indels: 2
DB: 29 Gaps: 1
US-10-029-020-14_COPY_1760_2300 (1-541) x AY413475 (1-6246)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db TACTACATCGGGCGCATGGCTCCTTGGGGTGGTGGCCACGGCATGGAGTGGCG 3267
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db CTGCAGACTGAGCCCACTTGTGGTGGCACCGTCAACCCACCGTGGGCAAGAGGAAT 3327
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAG 3387
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db GCTCGGGGCCAGGTCACTGTCTTTGGCGCGGCTGGGGTTTCACAACCGAAATCTCCTA 3447
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db TCTCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATGATGACCCACCGCAAGTTC 3507
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db ACCCTTCGGATTCTGTACGACCAGCGGGCGCGCCAGCCCTCTGGTCAACCCAGCAGCAGG 3567
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db CTGAATGGTGTCAACGTGACATACTCCCTCGGGGGTTACATTGCTGGCATCCAGAGGGGC 3627
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db ATCATGTCTGAAAGATGGAATACACGACGCGGGCGCGCATCACATCCAGGATCTTCGCT 3687
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLys-----SerMetValLeuLeuLeu 178
Db GATGGGAAGACATGGAGCTACACATACATCTTAGAAGAGGCAGGTTCCATGGTGTGCTACTA 3747

QY 179 HisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThr 198
Db CACAGCCAGAGCAGTATATCTTTGAGTTGACAAAGAATGACCGCCTCTTCTGTGACG 3807
QY 199 MetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsn 218
Db ATGCCCAACGTGGCGGCAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAAC 3867
QY 219 IleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHis 238
Db ATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGGCAC 3927
QY 239 LeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeu 258
Db CTCCTTTCACACCTTCTACCTGGGCACCTGGCCGAGGGTGATATACAAGTATGGCAAACTG 3987
QY 259 SerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThr 278
Db TCAAAGCTGGCAGAGACGCTCTATGACACACCACCAAGGTCAGTTTCACTATGACGAGACG 4047
QY 279 AlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyr 298
Db GCAGGCATGCTGAAGACCATCAACCTACAGAATGAGGGCTTCACTGCACCATCCGCTAC 4107
QY 299 ArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetVal 318
Db CGTCAGATTGGGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTC 4167
QY 319 AsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaVal 338
Db AACGCCCGTTTGTACTACAACTATGACAACAGCTTCCGGGTGACCCAGCATGCAGGCTGTG 4227
QY 339 IleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThr 358
Db ATCAACGAGACCCCACTGCCATTGATCTCTATCGCTATGATGATGTGTGAGGCAAGACA 4287
QY 359 GluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAla 378
Db GAGCAGTTTGGGAAGTTTGGTGTCTTTACTATGACATTAACAGATCATCACACAGCT 4347
QY 379 ValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGlu 398
Db GTCATGACCCACACCAAGCATTTTGTGATGCATATGGCAGGATGAAGGAAGTCAGTATGAG 4407
QY 399 IlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVal 418
Db ATCTTCCGCTCGCTCATGTACTGGATGACCGCAATCACTCGCTACTCTCTATGAGTATGAT 4467
QY 419 LysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAsp 438
Db AAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGAT 4527
QY 439 AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyr 458
Db GCTGACGGCCAGCTGCAGACAGTCTCCATCATGACAAGCCACTCTGGCGCTACAGCTAC 4587
QY 459 AspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeu 478
Db GACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTA 4647
QY 479 ArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGlu 498
Db CGGTATGACATCCGCAACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAG 4707
QY 499 AspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeu 518
Db GATGGCTTCTGAGGAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCCTGCTC 4767
QY 519 IleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGly 538
Db ATCAAGGCCTACAACCGGCTGGCAGCTGGAGTGTGAGTACCGTACCGTACGATGGCCTGGGG 4827
QY 539 ArgArgVal 541

Db 4828 CGGCGCGTG 4836

RESULT 2
AY413476

LOCUS
DEFINITION
Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION
AY413476

VERSION
AY413476.1 GI:39769438

KEYWORDS
SOURCE
ORGANISM
Pan troglodytes

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 5970)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL
Science 302 (5652), 1960-1963 (2003)

PUBMED
14671302

REFERENCE
AUTHORS
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE
Direct Submission

JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT
This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES
source
location/Qualifiers
1..5970
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5970
/locus_tag="HCM4903"

gene

ORIGIN

Alignment Scores:
Pred. No.: 2.27e-310 Length: 5970
Score: 2629.00 Matches: 505
Percent Similarity: 93.19% Conservative: 1
Best Local Similarity: 93.00% Mismatches: 35
Query Match: 92.41% Indels: 2
DB: 29 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x AY413476 (1-5970)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 2932 TACTACATCGGGCCGATGGCTCTTGGCGTGTCTGGCCAAACGGCATGGAGGTGGTG 2991

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 2992 CTGCAGACTGAGCCCCCACTTGTGGCTGGCACCGCTCAACCCACCCACCGTGGGCAAGAGGAAT 3051

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 3052 GTCACGCTGCCCATCGACACACGGCCTCAACCTGGTGAGTGGCGCCAGCGCAAGAGCAG 3111

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 3112 GCTCGGGGCCAGGTACTGTCTTTGGCGCGCGGTGCGGGTTCAACACCGAATCTCCTA 3171

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 3172 TCTCTGGACTTTGATCGCGTAAACACGCACAGAGAAGATCTATGATGACCCACCGCAAGTTC 3231

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 3232 ACCCTTCGGATTCTGTACGACAGCGGGGGCGCCAGCCTCTGGTCACCCAGCAGG 3291

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 3292 CTGAATGGTGTCAACGTGACATACTCCCTCGAGGTTACATTGCTGGCATCCAGAGGGGC 3351

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 3352 ATCATGTCTGAAAGAATGGATACGACCCAGGGGGCGGCATCATCCAGGATCTTCGCT 3411

QY 161 AspGlyLysThrTyrTipSerTyrThrTyrLeuGluLysSer-----MetValLeuLeuLeu 178
Db 3412 GATGGGAAGACATGGAGCTACACATACTTAGAAGAGCAGGCTNNNNNNNNNNNNNN 3471

QY 179 HisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThr 198
Db 3472 NNN 3531

QY 199 MetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsn 218
Db 3532 ATGCCCAACGTGGCGGCAGACNNNAGACCATCCGCTCAGTGGGCTACTACAGAAAC 3591

QY 219 IleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHis 238
Db 3592 ATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAGTCATACAGGACTTCACGTAGGATGGGCAC 3651

QY 239 LeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeu 258
Db 3652 CTCCTTCACACCTTCTACCTGGGCACCTGGCGCAGGGTGATATACAAAGTATGGCAAAC 3711

QY 259 SerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThr 278
Db 3712 TCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCAAGTTCACCTATGACGAGACG 3771

QY 279 AlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyr 298
Db 3772 GCAGGCTGTCTGAAGACCATCAACCTACAGAATGAGGGCTTCACCTGCACCATCCGCTAC 3831

QY 299 ArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetVal 318
Db 3832 CGTCAGATTGGGCCCCCTGATTGACCGACAGATCTTCCGCTTCACCGAGGAAGGCATGGTC 3891

QY 319 AsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaVal 338
Db 3892 AACGCCGCTTTTGACTACAACCTATGACAAACAGCTTCCGGGTGACCAAGCATGCAGGCTGTG 3951

QY 339 IleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThr 358
Db 3952 ATCAACGAGACCCCACTGCCCATGTGATCTCTATCGCTATGATGATGTGTAGGCAAGACA 4011

QY 359 GluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAla 378
Db 4012 GAGCAGTTTGGGAAGTTTGGTGTCAATTACTATGACATTAACCAAGATCATCACCACAGCT 4071

QY 379 ValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGlu 398
Db 4072 GTCATGACCCACACCAAGCATTTTGATGCATATGGCAGGATGAAGGAAGTGCAGTATGAG 4131

QY 399 IlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVal 418
Db 4132 ATCTTCCGCTCGCTCATGTACTGTGATGACTGTCCAGTATGATAACATGGGGCGAGTAGTG 4191

QY 419 LysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAsp 438
Db 4192 AAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCCTATGAGTATGAT 4251

QY 439 AlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyr 458
Db 4252 GCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTAC 4311

QY 459 AspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeu 478

Db 4312 GACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACACTA 4371
QY 479 ArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGlu 498
Db 4372 CGGTATGACATCCGGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAG 4431
QY 499 AspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeu 518
Db 4432 GATGGCTTNTGAGGCAGCGGCGGTGATATCTTTGAGTACAACTCANCTGNNTANN 4491
QY 519 IleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGly 538
Db 4492 ATCAAGNCNACACCGGCTGGCAGTGGAGTGTACGGTACCGCTACGATNNNTGGG 4551
QY 539 ArgArgVal 541
Db 4552 CGGCGGTG 4560

RESULT 3
AY405420 5094 bp DNA linear GSS 12-DEC-2003
LOCUS Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.

ACCESSION AY405420
VERSION AY405420.1 GI:39761394
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5094)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302
REFERENCE 2 (bases 1 to 5094)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.

FEATURES Location/Qualifiers
source 1..5094
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>5094
/locus_tag="HCM2218"

ORIGIN

Alignment Scores: 8.47e-262 Length: 5094
Pred. No.: 2233.00 Matches: 404
Score: 88.91% Conservative: 77
Percent Similarity: 74.68% Mismatches: 60
Best Local Similarity: 78.49% Indels: 0
Query Match: 29 Gaps: 0
DB:

US-10-029-020-14_COPY_1760_2300 (1-541) x AY405420 (1-5094)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db 2381 TACCAGATTGGTATGACGGCTCCCTCAGATTATCTACGCCAGTGGCCTGGACTCACAC 2440

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 2441 TACCAAACAGAGCCGCGACGTTCTGGCTGGCACCGCTAATCCGCGTTGCCAAAAGAAAC 2500
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 2501 ATGACTTTGCCTGGCGAGAACGGTCAAAACTTGGTGAATGGAGATTCCGAAAAGAGCAA 2560
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 2561 GCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTAAATGGCAGAAACCTCCTT 2620
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 2621 TCAGTTGACTTTGATCGAACAAACAAAGACAGAAAGATCTATGACCAACCCGTAAATTT 2680
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArg 120
Db 2681 CTACTGAGGATCGCTACGACACGCTGGGCACCCGACTCTCTGGTCCCAAGCAGCAAG 2740
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 2741 CTGATGGCGCTCAATGTACCTATTTCATCCACAGGTCAAATGGCCAGCATCCAGCGAGGC 2800
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 2801 ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTTGCT 2860
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 2861 GATGGTAAACATGAGGATTACACATATTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 2920
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 2921 CAGCGGCAGTACATCTTCGAATACGATATGTTGGGACCGCTGTCTGCCATCACCATGCC 2980
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 2981 AGTGTGGCTCGCCACACCATGCAGACCATCCGATCCATGGCTACTACCGCAACATATAC 3040
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 3041 AACCCCGGAAAGCAAGCGCTCCATCATCAGGACTACAACGAGGAAGGGTCTCTCTA 3100
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 3101 CAAACAGCTTTCTTGGGTACAAGTCGGAGGTCATTATCAATACAGAAAGGCAGACTAGG 3160
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 3161 CTCTCAGAAATTTTATATGATAGCACAAAGAGTCAGTTTACCTATGATGAAACAGCAGGA 3220
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 3221 GTCTAAAGACAGTAAACCTCCAGAGTGATGGTTTTTATTGCAACCATTAGATACAGGCAA 3280
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 3281 ATTGGTCCCCTGATTGACAGGCAGATTTCCTCGCTTTAGTGAAGATGGGATGGTAAATGCA 3340
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 3341 AGATTGACTATAGCTATGACAAACAGCTTTCGAGTGACCCAGCATGCAGGGTGTGATCAAT 3400
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 3401 GAAACGCCACTGCCTATTGATCTGTATCAGTTTGTATGACATTTCCTGGCAAAAGTTGAGCAG 3460
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrAlaValMet 380
Db 3461 TTTGGAAAGTTTGGAGTTATATATATATATTAACAGATCATTTCTACAGCTGTAAATG 3520
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400

```
Db 3521 ACCTATACGAAGCACCTTTGATGTCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC 3580
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 3581 AGGTCGCTCATGTACTAGATTACAATTCAAGTATGATAAACAATGGGTGGGTAAACCAAGAGA 3640
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 3641 GAGATTAAATAGGGCCCTTTTGCCAACACACCAAAATATGCTTATGAATATGATGTTGAT 3700
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTipArgTyrSerTyrAspLeu 460
Db 3701 GGACAGCTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACTACGATCTG 3760
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 3761 AATGGAACCTCCATTTACTGAACCCAGTAACAGTGGCGGTCTGACACCCCTTCGCTAT 3820
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 3821 GACCTGCGAGACAGAATCACTCGACITGGGTGAIGTTCAATATCGGTTGGATGAAGATGGT 3880
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 3881 TTCCTACGTCAAAGGGGCACGGAAATCTTTGAATATAGCTCCAAGGGGCTTCTAACTCGA 3940
QY 521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 3941 GTTTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCGTTATGACGGCCTGGGAAGCGGT 4000
QY 541 Val 541
Db 4001 GTT 4003

RESULT 4
AY405421 5069 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405421
VERSION AY405421.1 GI:39761395
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 5069)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5069)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..5069
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5069
/locus_tag="HCM2218"
gene
```

```
ORIGIN
Alignment Scores:
Pred. No.: 3.46e-261 Length: 5069
Score: 2228.00 Matches: 403
Percent Similarity: 88.72% Conservative: 77
Best local Similarity: 74.49% Mismatches: 61
Query Match: 78.31% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AY405421 (1-5069)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db 2381 TACCAGATTGGTTATGACGGCTCCCTCAGAATTATCTACGCCAGCGGCTGGACTCACAC 2440
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 2441 TACCAAAACAGAGCCGCACATTCTGGCTGGCACTGCTAATCCAAACCGTTGCCAAAGAAAC 2500
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
Db 2501 ATGACGTTGCCTGGTGAGAACCGTCAAAACTTGGTGGAAATGGAGATTCCGAAAAGAGCAA 2560
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 2561 GCCCAAGGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTT 2620
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 2621 TCAGTTGACTTTTGATCGAACAAACAAAGACAGAAAAGATCTATGACGACCACCCGTAATTT 2680
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 2681 CTACTGAGGATCGCCTACGACACGCTCTGGGCACCCGACTCTCTGGTCTGCCAAGCAGCAAG 2740
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 2741 CTGATGGCCGTCATGTCCACCTATTCATCCACAGGTCAAATGGCAGCATCCAGCGAGGC 2800
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 2801 ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTTTGTCT 2860
QY 161 AspGlyLysThrTrpSerTyrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180
Db 2861 GATGTTAAACATGGAGTTACATATTTAGAAAAGTCCCATGGTCTTCTGCTTCATAGC 2920
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 2921 CAGCGCAGTACATCTTCGAATACGATATGTGGGACCGCCTGTCTGCCATCACCATGCCCC 2980
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 2981 AGTGTGGCTCGCCACACCATCGACACCATCCGATCCATTGSGTACTACCGCAACATATAC 3040
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 3041 AACCCCCCGAAAGCAACGCCCTCCATCATCACGGACTACACGAGGAGGGGCTGCTTCTG 3100
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 3101 CAACACAGCTTCTTGGGTACAGTCGGAGGGTCTTATTCAAATACAGAAGGCAGACTAGG 3160
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 3161 CTCTCAGAAATTTTATATGATAGCACGAGAGTCAGTTTTTACCTACGATGAACACAGCAGGA 3220
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 3221 GTCCCTAAAGACAGTAAACCTCCAGAGTGACGGTTTTTATTGACCATTAGATACAGGCAA 3280
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
```


Db 3281 ATTGGTCCCCTGATTGACAGGCAGATTTTCGCTTTAGTGAAGATGGGATGGTAAATGCA 3340

Qy 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340

Db 3341 AGATTGACTATAGCTATGACAAACAGCTTTTCGAGTGACCAGCATGCAGGGTGTGATCAAT 3400

Qy 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360

Db 3401 GAAACTCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG 3460

Qy 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380

Db 3461 TTTGGAAAGTTTGGAGTTATATATTATGATATTAAACAGATCATTTCTACAGCTGTNATG 3520

Qy 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400

Db 3521 ACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATT 3580

Qy 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallyLys 420

Db 3581 AGGTCGCTCATGTACTGGATTACAAATTCAGTATGATAACATGGGTGGGTAAACCAAGAGA 3640

Qy 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440

Db 3641 GAGATTAATAATAGGGCCCTTTGCCAACACCAACCAATATGCTTATGAATATGATGTTGAT 3700

Qy 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460

Db 3701 GGACAGCTCCAAACAGTTTACCTCAATGAAAAAGATAATGTGGCGGTACAACTACGATCTG 3760

Qy 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480

Db 3761 AATGGAAACCTCCATTTACTGAACCCAACTAACAGTGCAGTGCAGTGCAGTGCAGTGCAT 3820

Qy 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500

Db 3821 GACCTCGGACAGACAGAATCACTCGACTGGGTGATGTTCAATATCGGTTGGATGAAGATGGT 3880

Qy 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520

Db 3881 TTCCTACGTCAAAGGGGCACGGAAATCTTGAATATAGCTCCAGGGGCTTCTAACTCGA 3940

Qy 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540

Db 3941 GTTTACAGTAAAGGCAGTGGCTGGACAGTGANNACCGTTTATGACGGCCTGGGAAGCGT 4000

Qy 541 Val 541

Db 4001 GTT 4003

RESULT 5

AY405422

LOCUS

DEFINITION

AY405422 5087 bp DNA linear GSS 12-DEC-2003

Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

ACCESSION

AY405422

VERSION

AY405422.1 GI:39761396

KEYWORDS

GSS.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 5087)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

AUTHORS

Infering nonneutral evolution from human-chimp-mouse orthologous gene trios

Science 302 (5652), 1960-1963 (2003)

PUBMED

14671302

REFERENCE

2 (bases 1 to 5087)

AUTHORS

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE

Direct Submission

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT

This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

Location/Qualifiers

source

1..5087

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

<1..>5087

/locus_tag="HCM2218"

gene

ORIGIN

Alignment Scores:

Pred. No.: 3.48e-261 Length: 5087

Score: 2228.00 Matches: 403

Percent Similarity: 88.72% Conservative: 77

Best Local Similarity: 74.49% Mismatches: 61

Query Match: 78.31% Indels: 0

DB: 29 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AY405422 (1-5087)

Qy 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db 2374 TACCAGATTGGGTATGATGGCTCCCTTAGAATCTTCTATGCCAGTGGTCTGGACTCTCAC 2433

Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 2434 TACCAGACAGAGCCCCACGTTCTGGCTGGCAGCGGAATCCCAAGTAGCCAAAGAAAC 2493

Qy 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60

Db 2494 ATGACTCTTCCCCGTGAGAACGGGCAGAAATCTGGTGGAGTGGAGATCCGAAAGAACAA 2553

Qy 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80

Db 2554 GCCCAGGGCAAAGTCAACGTATTTCGGCCGGAAGCTCAGGGTCAATGGCGCAACCTACTC 2613

Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100

Db 2614 TCAGTGGACTTTGATCGGACCACCAAGACGCGGAAAGATCTATGATGACCACCGGAAATTT 2673

Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120

Db 2674 CTCCTGAGGATCGCTTACGACACGTCGGGGCACCCGACTCTCTGGCTCCGAGTAGCAAG 2733

Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140

Db 2734 CTAATGGCAGTGAACGTCACCTACTCTATCCACCGGTCAAATTCAGCATCCAGAGAGGG 2793

Qy 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 2794 ACCACGAGCGAAAGGTGGACTATGACAGCCAGGGGAGGATCGTATCTCGGGTCTTTGCC 2853

Qy 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180

Db 2854 GATGGGAAACATGGAGTTACACGTACTTGGAAAAGTCCATGGTTCTTCTGCTCCATAGC 2913

Qy 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

Db 2914 CAGCGGCAGTACATCTTCGAATACGACATGTGGGACCGCCTGTCCGCCATCACCATGCCC 2973

Qy 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 2974 AGTGTGGCTCGCCACACCATGCAGACCATCCGGTCCATTGGCTACTACCGCAACATCTAC 3033

Qy 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 3034 AATCCCCAGAAAGCAATGCCTCTATCATCACCGACTACAACGAGGAAGGGCTGCTTCTG 3093
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 3094 CAAACAGCTTTCCTGGGAACGAGTCGGAGGGTCTTATTCAAGTATAGAAGGCAGACACGAG 3153
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 3154 CTATCAGAAATTTTATACGACAGCACAGAGTCAGTTTACCTACGACGAACAGCGGGA 3213
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 3214 GTCCTGAAACACAGTAAACCTTCAGAGTGTGGTTTATTTCACCATTTAGATACAGGCAA 3273
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 3274 ATTGGTCCCTGATTGACAGACAGATTTTCCGCTTCAGCGAGGATGGATGGTAAATGCG 3333
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 3334 AGATTTGACTATAGCTACGACACACAGCTTTCGAGTGACCAGCATGCAGGTGTCATCAAT 3393
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 3394 GAAACACCACTGCCCATTTGATCTATACCAAGTTTGATGACATCTCTGGCAAAAGTCGAGCAG 3453
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 3454 TTTGGAAAATTGGAGTGATATATACTACGACATCAACCAATCATTTCCACGGCCGTGATG 3513
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 3514 ACTTATACAAAGCACTTTGATGCTCATGGCGGCATCAAGGAGATCCCAATATGAGATATT 3573
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 3574 AGGTCACTCATGTACTGGATTACAAATTCAATATGATAATATGGCCGGGTAAACCAAGAGA 3633
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 3634 GAGATTAAATTTGGCCCTTTTGCCCACTACCAATACCGGTACGAGTACGACGTCGAT 3693
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db 3694 GGACAGCTCCAACACAGTTTACCTAAACGAAAGATCATGTGGCGGTACAACCTACGACCTA 3753
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 3754 AATGGAACCTCCACTTGCTCAACCCAGCAGCAGCGCCCGCTGACCCCTCTGCGCTAT 3813
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 3814 GACCTGCGCGACAGAAATCACCCGCTGGCGGATGTTTCAGTACCGGCTGGATGAAGATGGT 3873
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 3874 TTCCTGCGTCAGAGGGGCACTGAAATTTTGAATACAGCTCCAAGGGCTTCTGACTCGA 3933
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 3934 GTCTACAGTAAAGGCAGTGGCTGGACAGTGATCTATCGGTACGACGGCCTGGGAAGACGT 3993
QY 541 Val 541
Db 3994 GTT 3996

RESULT 6
AK037897

LOCUS Mus musculus 16 days neonate thymus cDNA, RIKEN full-length
DEFINITION enriched library, clone:A130060L24 product:odd Oz/ten-m homolog 1
(Drosophila), full insert sequence.
3038 bp mRNA linear HTC 19-SEP-2003

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK037897
AK037897.1 GI:26332325
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3038)

ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,
FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W.,
HAYASHIDA, K., HAYATSU, N., HIRAMOTO, K., HIRAOKA, T., HIROZANE, T.,
HORI, F., IMOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T.,
KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KOUDA, M.,
KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M.,
NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHSAO, N.,
OKAZAKI, Y., SAITO, R., SAITO, H., SAKAI, C., SAKAI, K., SAKAZUME, N.,
SANO, H., SASAKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T.,
SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKAHIRA, S.,
TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A.,
MURAMATSU, M. and HAYASHIZAKI, Y.
Direct Submission

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/

COMMENT

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.
FEATURES Location/Qualifiers
source 1.3038
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:Al30060L24"
/db_xref="MGI:2402337"
/db_xref="taxon:10090"
/clone="Al30060L24"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
CDS
<1..2978
/note="unnamed protein product; odd Oz/ten-m homolog 1 (Drosophila) (MGD|MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%, match=3037)
putative"
/codon_start=3
/protein_id="BAC29893.1"
/db_xref="GI:26332326"
/translation="EINLSSEPHILAGAVNP TLGKNISLPGEHNANLIEWRQKQON KGNVAFERRLRRAHNRNLLSIDFDHMTRTGKIYDHRKFTLRILYDQTRPILWSPVS RYNEVNITYSPGLVTFIORGTWNEKMEYDQSGKIISRTWADGKIWSYTYLEKSVMLL LHSQRRYIFEYDQSDCLSVTPMSVRHSLQTM LSVGYRNIYTPPDSSTSFIDYSR DGRLLQTLHLGTGRRLVYKTKQARLSEILYDTTQVTLTYEESGVIKTIHLMHDGFI CTIRYROTGLRIQRIFRSEEGLVNARFDYSNNFRVTSMQAVINETPLPIDLYRVV DVSGRTEQFGKFSVINYLNQVITTTVMKHTKIFNANGQVIEVQYEILKAIAYWMTIQ YDNMGRMVICDIRVGVDANITRYFYEYDADGQLOQTVSVNDKIQWRYSYDLNGNINLLS HGNSARLTPLRYDLRLDRITLGEIQKMDDEGFLRQGRNDIFEYNSNGLLQKAYNKVS GWTQVYYDGLGRRVASKSSGLQHLQFFYADLANPIRVTHLYNHTSAEITSLYLDQG HLIAMELSSGEEYYVACDNMGTPLAVFSRRGQVKEILYTPYGDYHDTYDPDEFVILG FHGGLYDFLTKLHLGQRDYDVVAGRWTTTPNHIIWKQLNLLPKPFFNLYSFENNYPVGK IQDVAKYTTDITGWLFLGFLQHLNVLPGPKPELENMELTYELLQLOTKTQEWDPGKM ILGIQCELOKQLRNFISLDQLEPMTPOYNEGRCLGKGKQPRFAAVPSVFGKIGKFAIKE ILGADIIIGVANEDSRRLAAILNNAHYLENLHFTIEGRDTHYFIKLSLEEDLVILGN TGGRRILEGNVNTVSQMTSVLNGRTRRFADIQLOHGALCFNIRYGTITVEEEKHNVLE MARQRAVAQAWTQEQRRLOEGEGEGRVWTEGEKQQLLGTGRVQGYDYFVLSVEQYLE LSDSANNIHFMRQSEIGRR"
ORIGIN
Alignment Scores: 1.08e-235 Length: 3038
Pred. No.: 2017.50 Matches: 371
Score: 85.50% Conservative: 77
Percent Similarity: 70.80% Mismatches: 75
Best Local Similarity: 70.91% Indels: 1
Query Match: 11 Gaps: 1
DB: 1
US-10-029-020-14_COPY_1760_2300 (1-541) x AK037897 (1-3038)
QY 18 GluValAlaLeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGly 37
Db 3 GAGATCAACCTCAGCTCGAGCGCCACATCCTGGCAGGGGCTGTCAACCCCACTTAGGC 62
QY 38 LysArgAsnValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArg 57
Db 63 AAATGCAACATATCGTTGCCTGGAGAGCACAAATGCAAAACCTCATTGAGTGGAGACAGAGA 122
QY 58 LysGluGlnAlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArg 77
Db 123 AAGGAGCAAAACAAAGGCAACGTTTCAGCTTTTGAGAGAAAGACTGAGGGGCCATAACAGA 182
QY 78 AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspHis 97
Db 183 AACCTACTTTCATTGATTTTGTATCATATGACCCGCGCGGTAAAGATCTATGATGACCAC 242
QY 98 ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPro 117
Db 243 CGAAATTCACCTCTCGAATTCCTTATGACCAGACCGGGCGACCTATTCTGTGTCCTCT 302
QY 118 SerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIle 137

Db 303 GTGAGCAGATATAATGAAGTGAAACATCACGTTATTCACCTTCAGGATTAGTCACATTTATT 362
QY 138 GlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArg 157
Db 363 CAAAGAGGAACCTGGAATGAAAGATGGAGTATGACAGAGTGGAAAAATAATCTCAAGA 422
QY 158 IlePheAlaAspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeu 177
Db 423 ACTTGGCCGATGGGAAATTTGGAGCTATACCTACTATAGAAAAGTCTGTGATGCTTCTG 482
QY 178 LeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerVal 197
Db 483 TTACACAGTCAGCGCGGTACATCTTTGAGTATGACCAATCAGATTGCCTGCTGTCAGTT 542
QY 198 ThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArg 217
Db 543 ACCATGCTAGCATGGTGGCGGCACAGCTTACAAACCATGCTTTTTCAGTGGGCTACTACCT 602
QY 218 AsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGly 237
Db 603 AATATCTACACCCACGACAGACAGCAGCACCTCTTTTATCCAGACTACAGTCGAGATGCG 662
QY 238 HisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLys 257
Db 663 CGACTGTTACAGACTCTGCATCTAGGACTGGGCGCAGAGTTTATACAGTACACAAAG 722
QY 258 LeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGlu 277
Db 723 CAAGCAAGACTGTCTGAGATTCTCTATGACACCACCTCAGGTCACATTAACATATGAAGAG 782
QY 278 ThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArg 297
Db 783 TCTTCTGGAGTAATTAAGACAATAACACTGATGATGATGATGATGATGATGATGATGAT 842
QY 298 TyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMet 317
Db 843 TACAGGCAACAGGACCTCTCATTTGGACGTCAGATTTTCCGATTCTGAGTGAAGAGGTCCT 902
QY 318 ValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAla 337
Db 903 GTGAATGACGGTTTGACTACAGTTAC---AACAACTTTCGTGTCCAGCATGCAAGCT 959
QY 338 ValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLys 357
Db 960 GTGATCAATGAGACCCCTTTTGCCCATAGATCTGATACCGATATGTTGATGTTTCTGGTAGA 1019
QY 358 ThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThr 377
Db 1020 ACAGAACAGTTTGGAAAAATTTCAGTGTAAATTAATGATTTAAATCAGGTTATAACTACA 1079
QY 378 AlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyr 397
Db 1080 ACAGTGATGAAGCACACCAAGATTTTCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1139
QY 398 GluIlePheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgVal 417
Db 1140 GAGATCCTAAAGGCAATGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 1199
QY 418 ValLysLysGluLeuLysValGlyProTyrTyrAlaAsnThrThrArgTyrSerTyrGluTyr 437
Db 1200 GTCATATGTGATATAAGGTAGGAGTAGATGCCAATATATCACAAGGTACTTCTATGATAT 1259
QY 438 AspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSer 457
Db 1260 GATGCTGATGGGCAACTTCAAACTGTTTCTGTGATGATGATGATGATGATGATGATGAT 1319
QY 458 TyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrPro 477
Db 1320 TATGATCTGAATGGAAACATCAACCTTTTGGCCCATGGAATAGTGTCTGCTTACTCCT 1379
QY 478 LeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAsp 497
Db 1380 CTCCGATATGACCTCCGAGACCGCATCACTAGACTAGGGGAAATTCAGTATAAAATGGAT 1439

QY 498 GluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeu 517
Db 1440 GAAGACGGTTTCTGACACAGAGGGGAAATGATATATTGAGTATAATCTAATGGTCTG 1499

QY 518 LeuIleLysAlaTyrAsnArgAlaGlySerTyrPheValArgTyrArgTyrAspGlyLeu 537
Db 1500 CTGCAGAAAGCCTACAAATAAGTTTCTGGCTGGACTGTACAGTATTACTATGATGGGCTT 1559

QY 538 GlyArgArgVal 541
Db 1560 GGGCGACGCGTT 1571

RESULT 7
BU151768
LOCUS
DEFINITION
AGENCOURT_8750792 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335962
5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 894)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.

cDNA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLA13799 row: d column: 11
High quality sequence stop: 677.
Location/Qualifiers
1..894

FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6335962"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_130"
/note="Organ: otocysts; Vector: pCMV-SPORT6.1; Site_1:
EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.95 kb. Constructed by
ResGen, Invitrogen Corp. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
Pred. No.: 2.87e-133 Length: 894
Score: 1177.00 Matches: 208
Percent Similarity: 93.51% Conservative: 37
Best Local Similarity: 79.39% Mismatches: 17
Query Match: 41.37% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x BU151768 (1-894)

QY 280 GlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArg 299
Db 3 GGAGTCCTGAAACACAGTAAACCTTCAGAGTGATGGTTTATTTCACCATTAGATACAGG 62

QY 300 GlnIleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsn 319
Db 63 CAAATTGGTCCCTGATTGACAGACAGATTTTCCGCTTCAGCGAGGATGGAATGGTAAAT 122

QY 320 AlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db 123 GCGAGATTGTACTATAGCTACGACAACAGCITTCGAGTGACCATGCAGGGTGTATC 182

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359
Db 183 AATGAAACACCATGCCCATTGATCTATACCAGTTGATGACATCTCTGGCAAAGTCGAG 242

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrAlaVal 379
Db 243 CAGTTTGGAAAATTGGAGTGATATACTACGACATCAACCAATCATTTCCACGCCGTG 302

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 303 ATGACTTATACAAAGCACTTTGATGCTCATGGCGCATCAAGGAGATCCAATATGAGATA 362

QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLys 419
Db 363 TTTAGTCACTCATGTACTGGATTACAATTCAATATGATAATATGGCCGGGTAACCAAG 422

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 423 AGAGAGATTAAATTTGGCCCTTTTGCCAACTACCAATACGCGTACGAGTACGACGTC 482

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
Db 483 GATGGACAGCTCCAAACAGTTTACCTAAACGAAAGATCATGTGGCGGTACAACTACGAC 542

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 543 CTAATGGAACCTCCACTTGCTCAACCCAGCAGCAGCGCCCGCTGACCCCTCTGCGC 602

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 603 TATGACCTGGCGACAGAAATCACCCGCTGGCGATGTTTACGTACCGCTGGATGAAGAT 662

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuIle 519
Db 663 GGTTCCTGCGTCAGAGGGGCACTGAAATTTTGAATACAGCTCCAAAGGGCTTCTGACT 722

QY 520 LysAlaTyrAsnArgAlaGlySerTyrPheValArgTyrArgTyrAspGlyLeuArg 539
Db 723 CGAGTCTACAGTANAGGCAGTGGCTGGACAGTGATCTATCGTACGACGCGCTGGGAAGA 782

QY 540 ArgVal 541
Db 783 CGTGT 788

RESULT 8
BU610605
LOCUS
DEFINITION
UI-M-FC0-caq-f-07-0-UI.r1 NIH BMAP_FCO Mus musculus cDNA clone
UI-M-FC0-caq-f-07-0-UI 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, Univeristy of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project

BU610605 834 bp mRNA linear EST 20-FEB-2003
UI-M-FC0-caq-f-07-0-UI.r1 NIH BMAP_FCO Mus musculus cDNA clone
UI-M-FC0-caq-f-07-0-UI 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 834)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, Univeristy of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
This clone was contributed by the Brain Molecular Anatomy Project

Db	23	AGCGAARAGGTGGACTATGACAGCGAGGGAGGATCGTATCTCGGGTCTTTGCCGATGGG	82
QY	163	LysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSerGlnArg	182
Db	83	AAACATGGAGTTACAGTACTTGGAAAAGTCCATGGTTCTTCTGCTCCATAGCCAGCGG	142
QY	183	GlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetProAsnVal	202
Db	143	CAGTACATCTTCGATACGACATGTGGGACCGCTGTCCGCCATCACCATGCCAGTGTG	202
QY	203	AlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnPro	222
Db	203	GCTCGCCACACCATGCGAGCCATCCGGTCCATTGGCTACTACCGCAACATCTACATCCC	262
QY	223	ProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThr	242
Db	263	CCAGAAAGCAATGCCTCTATCATCACCGACTACAACGAGGAAGGGTGCTTCTGCAACA	322
QY	243	PheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLysLeuAla	262
Db	323	GCTTTCCTGGGAACGAGTTCGAGGGTCTTATTCAAGTATAGAAGGCAGACCAGGCTATCA	382
QY	263	GluThrLeuTyrAspThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeu	282
Db	383	GAAATTTTATACGACAGCAAGAGTCAGTTTACTACGACGAAACAGCGGGAGTCCTG	442
QY	283	LysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGly	302
Db	443	AAACAGTAACCTTCAGAGTGATGGTTTATTTCACCATTAGATACAGCAAAATGGT	502
QY	303	ProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAlaArgPhe	322
Db	503	CCCCTGATTGACAGACAGATTTCCCGCTTCAGCAGGATGGAATGGTAAATCGGAGATT	562
QY	323	AspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluThr	342
Db	563	GACTATAGCTACGACAACAGCTTTCGAGTGACCAGCATGCAGGGTGTCAATCAATGAA	622
QY	343	ProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGly	362
Db	623	CCACTGCCCATGATCTATACCAAGTTTGATGACATCTCTGGCAAGTCGAGCAGTTTGA	682
QY	363	LysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThr-AlaValMetThrHi	382
Db	683	AAATTCGGAGTGATATACTACGACATCAACCAAAATCATTTCCACGGGNCGTGATGAC	742
QY	382	sThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyr-GluIlePheArgS	402
Db	743	TACAAAGCACTTTGTATGCTCATGGGCGCATCAAGGAGATCCCAATATGGAGATATT	802
QY	402	erLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArg---ValValLysLysG	421
Db	803	CACATCGGACTGGATTACAATTCATATGATAATATGGCGCGGGTAACCAAGAGAGAG	862
QY	421	luLeuLysValGlyProTyrAlaAsn-ThrThrArgTyrSer	434
Db	863	AATTAAATTGGGGCTTTTGGCAACAACACTACCAAAATACGCG	904
RESULT 10			
CF723353			
LOCUS			
DEFINITION			
UI-M-GV0-cjh-j-04-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone			
IMAGE:30546171 5', mRNA sequence.			
CF723353			
CF723353.1 GI:37597521			
EST.			
SOURCE			
Mus musculus (house mouse)			
ORGANISM			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 694)			
REFERENCE			
AUTHORS			
NIH-MGC http://mgc.nci.nih.gov/ .			
TITLE			
National Institutes of Health, Mammalian Gene Collection (MGC)			

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

FEATURES

Location/Qualifiers
1..694
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30546171"
/tissue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GV0"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAACTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 3.72e-129 Length: 694
Score: 1142.00 Matches: 225
Percent Similarity: 98.27% Conservative: 2
Best Local Similarity: 97.40% Mismatches: 3
Query Match: 40.14% Indels: 2
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x CF723353 (1-694)

QY	78	AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHis	97
Db	4	AACCTCTTGTCTTTGGACTTTGACCGTGTACACGACAGAGAAGATCTACGATGACCAT	63
QY	98	ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerPro	117
Db	64	CGCAAGTTCAACCTTCGGATCCTATATGACCAGGAGGAGGCCACCTCTGGTCACCT	123
QY	118	SerSerArgLeuAsnGlyValAsnValThrTyrSerPro-GlyGlyTyrIleAlaGlyIle	137
Db	124	AGCAGCAGGCTGAATGGTGTAAATGTGACCTACTCCCTGGGAGGTTCACATTGCTGGAAT	183
QY	137	eGlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerAr	157
Db	184	CCAAAGGGGCATCATGTCTGAGAGAATGGATATGATCAGGCGGCGCATCACATCCCG	243
QY	157	gilePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLe	177
Db	244	GATCTTTGCAGACGGGAAAATGTGGAGCTACACGTACTTAGAGAAGTCCATGGTCTTCA	303
QY	177	uLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerVa	197
Db	304	TCTCCACAGCCAGGAGGCAGTACATCTTCAGTTTGACAAGAATGACCGCCTCTCTTCTGT	363

is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores: 1.84e-125 Length: 758
Pred. No.: 1112.50 Matches: 194
Score: 93.20% Conservative: 39
Percent Similarity: 77.60% Mismatches: 16
Best Local Similarity: 39.10% Indels: 1
Query Match: 14 Gaps: 1
DB:

US-10-029-020-14_COPY_1760_2300 (1-541) x CB723840 (1-758)

QY 291 GlyPheThrCysThrIleArgTyrArgGlnIleGlyProIleAspArgGlnIlePhe 310
Db 9 GGTTCCTCTGTACCATCAGGTACCGAAAGTTGGGCCCTTGTGGACAAGCAGATTTAC 68
QY 311 ArgPheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsnTyr---AspAsnSer 329
Db 69 AGTTCTCTGAGGAAGGAATGATCAACGCCAGGTTTGATTATACCTATCAGCAATAGC 128
QY 330 PheArgValThrSerMetGlnAlaValIleAsnGluThrProIleAspLeuTyr 349
Db 129 TTCCGCATGCCAGCATCAACCCGTCATTAGCGAGACTCCCTTCTCTGTGACCTCTAC 188
QY 350 ArgTyrAspAspValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyr 369
Db 189 CGCTATGATGAGATTTCCGGCAAGGTGGAACACATTCGGCAAGTTTGGGTCTACTACTAC 248
QY 370 AspIleAsnGlnIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyr 389
Db 249 GACATCAACCATGATCATCACCATGCTCCGTCATGACGCTTAGCAAGCACCTTGACACCCAT 308
QY 390 GlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrVal 409
Db 309 GGGCGCATCAAGGAAGTGAATATGAGATGTTCCGGTCCCTCATGTACTGGTACTGTG 368
QY 410 GlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsn 429
Db 369 CAATATGACAGTATGGGTAGGTGTCATCAAGAGGGAACCTAGGCCCCCTATGCCAAC 428
QY 430 ThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsn 449
Db 429 ACCACAAAGTACACCTATGACTATGACGGGACGGGACGGGACGCTCCAGAGTGTGGCCGTCAT 488
QY 450 AspLysProLeuTyrArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerPro 469
Db 489 GACCGGCTACCTGGCGCTATAGCTATGACCTCAATGGGAACCTGCACCTTCTAAACCCA 548
QY 470 GlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeu 489
Db 549 GGAAACAGTGTCTGCCTCATGCCCTTACGCTATGACCTCCGTGACCGGATAACAGGCTA 608
QY 490 GlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIle 509
Db 609 GGGACGTGCGAGTACAAATCGATGACGATGGCTATTTGTCCAGAGAGGGTCCAGACATC 668
QY 510 PheGluTyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSer 529
Db 669 TTTGAATACAACTCCAAGGGCTCTCTGACAAGAGCATACAAAGGCCCGGATGGAGC 728
QY 530 ValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 729 GTGCACTACCGCTATGACGGAGTGGGCGCG 758

RESULT 12

BE741784

LOCUS

DEFINITION

BE741784 757 bp mRNA linear EST 15-SEP-2000
LOCUS 601595546F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949571 5',
DEFINITION

QY 197 lThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrAr 217
Db 364 GACCATGCCAACGTAGCCCGCAGACGCTGGAGACCATCCGCTCAGTGGGCTACTACAG 423
QY 217 gAsnIleTyrGlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGl 237
Db 424 GAACATCTACAGCCCGCCGGAAGCAACGCCTCAGTCTCAGTCTCAGGACTTCACCTGAGGATGG 483
QY 237 yHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLy 257
Db 484 ACACCTGTTACATACCTTCTACTTGGCACCGGCCCGGGTGATTTACAAGTATGGCAA 543
QY 257 sLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGl 277
Db 544 GTTGTCAAAGCTGGCCGAGACTCTGTATGACACCACTAAGGTAAAGCTTCACCTACGACGA 603
QY 277 uThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleAr 297
Db 604 GACCGCAGGATGCTGAAGACTGTCAACCTACAGAAATGA-GGGTTCACCTGCATATCCG 662
QY 297 gTyrArgGlnIleGlyProLeuIleAspArg 307
Db 663 CTACCGTCAGATTGGGCCCTGATTGATCGG 693

RESULT 11

CB723840

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB723840 758 bp mRNA linear EST 09-JUL-2003
UI-M-FY0-cez-f-24-0-UI.x1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6844153 5', mRNA sequence.

CB723840 GI:29780982

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 758)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at

<http://genome.uiowa.edu/distribution/mousefl.html>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1. .758

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6844153"

/tissue_type="whole brain"

/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_FY0"

/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;

Site_2: Not I; The library was constructed according

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with EcoR I adaptor, digested with NotI and then cloned

directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

ACCESSION BE7411784 mRNA sequence.
VERSION BE7411784.1 GI:10155776
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM814 row: k column: 12
High quality sequence start: 24
High quality sequence stop: 755.
Features
Location/Qualifiers
1..757
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3949571"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
Alignment Scores:
Pred. No.: 2.7e-124 Length: 757
Score: 1103.00 Matches: 199
Percent Similarity: 95.87% Conservative: 33
Best Local Similarity: 82.23% Mismatches: 10
Query Match: 38.77% Indels: 1
DB: 10 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x BE7411784 (1-757)
QY 263 GluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeu 282
Db 22 GAGAAATTATATGATAGCACAAAGAGTCAGTTTACCTATGATGAAACAGCAGGAGTCCTA 81
QY 283 LysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGly 302
Db 82 AAGACAGTAAACCTCCAGAGTGATGGTTTATTGTCACCATGATACAGGCAAT-GGT 140
QY 303 ProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAlaArgPhe 322
Db 141 CCCTGATTGACAGGCAGATTTCGCTTTAGTGAAGATGGATGGTAAATGCAAGATT 200
QY 323 AspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluThr 342
Db 201 GACTATAGCTATGACAAACAGTTTCGAGTGACCAGCATCGAGGTGTGATCAATGAAACG 260
QY 343 ProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGly 362
Db 261 CCACTGCCTATTGATCTGTATCAGTTTGTATGACATTCTGGCAAAGTTGAGCAGTTTGA 320
QY 363 LysPheGlyValIleTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHis 382

Db 321 AAGTTTGGAGTTATATATTATGATATTAAACCATCATTTCTACAGCTGTAATGACCTAT 380
QY 383 ThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSer 402
Db 381 ACGAAGCACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTAGGTTCG 440
QY 403 LeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeu 422
Db 441 CTCATGTACTGGATTACAATTCAATGATGATAACATGGGTCGGGTAAACCAAGAGAGATT 500
QY 423 LysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGln 442
Db 501 AAATAGGGCCCTTGCCCAACACCACCAATATGCTTATGAATATGATGTGATGGACAG 560
QY 443 LeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeuAsnGly 462
Db 561 CTCCAAACAGTTTACCTCAATGAAAAGATAATGTGGCGGTACAACCTACGATCTGAATGGA 620
QY 463 AsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIle 482
Db 621 AACCTCCATTACTGAACCCCAAGTAACATGCGCGTCTGACACCCCTTCGCTATGACCTG 680
QY 483 ArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeu 502
Db 681 CGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGATGGTTTCCTA 740
QY 503 ArgGln 504
Db 741 CGTCAA 746

RESULT 13
CNS03W3H
LOCUS
DEFINITION
CNS03W3H 868 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone 063M14 of library G from Tetraodon nigroviridis, genomic survey sequence.
AL263222
AL263222.1 GI:7984874
GSS: genome survey sequence.
Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidea; Tetraodontidae; Tetraodon.
1
Roest Croliius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
10835645
2
Roest Croliius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10899143
3 (bases 1 to 868)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at

263 GluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeu 282
Db 22 GAGAAATTATATGATAGCACAAAGAGTCAGTTTACCTATGATGAAACAGCAGGAGTCCTA 81
QY 283 LysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGly 302
Db 82 AAGACAGTAAACCTCCAGAGTGATGGTTTATTGTCACCATGATACAGGCAAT-GGT 140
QY 303 ProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAlaArgPhe 322
Db 141 CCCTGATTGACAGGCAGATTTCGCTTTAGTGAAGATGGATGGTAAATGCAAGATT 200
QY 323 AspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluThr 342
Db 201 GACTATAGCTATGACAAACAGTTTCGAGTGACCAGCATCGAGGTGTGATCAATGAAACG 260
QY 343 ProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGly 362
Db 261 CCACTGCCTATTGATCTGTATCAGTTTGTATGACATTCTGGCAAAGTTGAGCAGTTTGA 320
QY 363 LysPheGlyValIleTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHis 382

http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
Location/Qualifiers
1. .856
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="063M14"
/clone_lib="G"
/note="Genoscope sequence ID : COBG063BG07SP1-end :
PUC-Ori"
ORIGIN
Alignment Scores:
Pred. No.: 1.06e-123 Length: 868
Score: 1099.00 Matches: 201
Percent Similarity: 90.37% Conservative: 43
Best Local Similarity: 74.44% Mismatches: 26
Query Match: 38.63% Indels: 1
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x CNS03W3H (1-868)
QY 172 LysSerMetValLeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsn 191
Db 58 CAGTCAATGGTCCTCCTGCTGCACAGTCAACGGCAGTACATATTTGACTCTCAG 117
QY 192 AspArgLeuSerSerValThrMetProAsnValAlaAargGlnThrLeuGluThrIleArg 211
Db 118 GACCGGTGGCTGCCGTCACAATGCCAGTGTGGCCGTTACACCATGCAGACGATCCGC 177
QY 212 SerValGlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGln 231
Db 178 TCGGTGGCYMCYMCAGAAACCTTTACACCCCTCCAGAAAGCAACGCTCGGTGGCGTA 237
QY 232 AspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgVal 251
Db 238 GATTACAGCAGGATGGCCTCCTGCTGAGAGTGGCTCACTTGGGCACGGTTCGCCGGTC 297
QY 252 IleTyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysVal 271
Db 298 CTCTACAGTACCGCCGGCAGAACAAAGCTGTGCGAGATCCTGTACGACAGCACCAGGGTC 357
QY 272 SerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGly 291
Db 358 AGTTTTACCTACGACGAGCGCGCAGCTTGATGAAGACCGTTAACCTGCAGAAATGAGGC 417
QY 292 PheThrCysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArg 311
Db 418 TTCATCTGCTCCATTGTTATCGCCAAAGTGGGCCCACTGATCGACAGGAGATTTCCGC 477
QY 312 PheThrGluGlyMetValAsnAlaAargPheAspTyrAsnTyrAspAsnSerPheArg 331
Db 478 TTCAGCGAAGACGGCATGGTCAATGCGCGCTTTCATTACACCTACGACAGCAGCCTGCGT 537
QY 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351
Db 538 GTCACCAGCGTGAAGCGGTCAATGAACACCACTGCCAATCGATCTCTATCAGTAT 597
QY 352 AspAspValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIle 371
Db 598 GATGACATCTCGGGAAGGTGGAACAGTTTGGAAGTTTGGAGTCATCTATTATGATATC 657
QY 372 AsnGlnIleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArg 391
Db 658 AATCAATCATATCCCGCAGTTATGACCTACACCAACACATTTGATGCACAGGACGG 717
QY 392 MetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyr 411
Db 718 ATCAAGGAGATCCAGTATGAGATATTCGTTTCGCTCATGTACTGGATCACCTTCCAGTAT 777
QY 412 AspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThr 431
Db 778 GACGATGTGGGGCGAGTGAAGAGAGAGATCAAAATC-GGCCCCCTTTGCGGAACACCACC 836

QY 432 ArgTyrSerTyrGluTyrAspAlaAspGly 441
Db 837 AAGTACGCTWACGAGTACGACGTGGACGGC 866
RESULT 14
CF534506
LOCUS
DEFINITION
IMAGE:30536758 5', mRNA sequence.
ACCESSION
CF534506
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 856)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1. .856
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30536758"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_G10"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 1.82e-123 Length: 856
Score: 1097.00 Matches: 205
Percent Similarity: 98.57% Conservative: 2
Best Local Similarity: 97.62% Mismatches: 3
Query Match: 38.56% Indels: 0
DB: 14 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x CF534506 (1-856)
QY 332 ValThrSerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyr 351
Db 3 GTGACTAGCATGCAGGCTGTGATCAATGAGACCCCACTGCCACCTCTACCGCTAT 62

QY 352 AspAspValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIle 371
Db 63 GATGATGTGTcAGGGAAGACAGACAGATTtGGGAAGTTtGGTGTcATCTACTACGACATC 122

QY 372 AsnGlnIleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArg 391
Db 123 AACCAGATcATTACCACAGCGGTcATGACCCACACCAAGCACTtTGATGCTtTATGGCAGG 182

QY 392 MetLysGluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyr 411
Db 183 ATGAAGGAAGTACAGTATGAGATtTtCCGGTCACTcATGTACTGGATGACTGtTCAGTAT 242

QY 412 AspAsnMetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThr 431
Db 243 GATAACATGGGACGGGTAGTGAAGAAGGAGCTGAAGGTGGGACCTATGCCAACACTACC 302

QY 432 ArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLys 451
Db 303 CGCTACTCTCTATGAGTATGATGTCTGATGGCCAGCTGCAGACAGTCTCCATCAATGACAAG 362

QY 452 ProLeuTrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsn 471
Db 363 CCACtCTGGCGCTACAGCTATGACCTCAATGGGAACCTACACTtTGCTGAGCCCTGGGAAC 422

QY 472 SerAlaArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAsp 491
Db 423 AGCGACGGCTCACACCACtACTACGGTATGACCTCCGTGACCGCATCACTAGGCTGGGTGAT 482

QY 492 ValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGlu 511
Db 483 GTACAGTACAGATGGATGAGGATGGCTtCCCTGAGGCAGCGGGTGGGATGTCTTCGAG 542

QY 512 TyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArg 531
Db 543 TACAATTcAGCGGGCTGCTCATCANAGCCTACAACCGGGCTAGTGGGTGGAGTGTcAGG 602

QY 532 TyrArgTyrAspGlyLeuGlyArgArgVal 541
Db 603 TACCGATACGATGGCTGGGACGCCGAGTA 632

RESULT 15
CNS02PHB
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of clone
155F06 of library G from Tetraodon nigroviridis, genomic survey
sequence.
AL207992
VERSION
AL207992.1 GI:7866811
KEYWORDS
GSS: genome survey sequence.
Tetraodon nigroviridis
SOURCE
Tetraodon nigroviridis
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.
1
Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
MEDLINE
PUBMED
10835645
REFERENCE
2
Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837

PUBMED 10899143
REFERENCE 3 (bases 1 to 1031)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Location/Qualifiers
source
1..1031
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="155F06"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG155DC03LPI-end : T7"

ORIGIN

Alignment Scores:
Pred. No.: 9.27e-120 Length: 1031
Score: 1068.00 Matches: 202
Percent Similarity: 92.98% Conservative: 10
Best Local Similarity: 88.60% Mismatches: 16
Query Match: 37.54% Indels: 1
DB: 29 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x CNS02PHB (1-1031)

QY 314 GluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThr 333
Db 11 GAAGAGGGGATGGTCAC-GCCAGATTAGACTATGTCTATGACAACAGTTTCAGAGTTACA 69

QY 334 SerMetGlnAlaValIleAsnGlnThrProLeuProIleAspLeuTyrArgTyrAspAsp 353
Db 70 AGCATGTAGCBGTGATCAATGAACCTCCACTGCCCATCGACTTGTATCGCTACGACGAC 129

QY 354 ValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGln 373
Db 130 GTGTCtGGCAAAACAGAGCAGtTTGGTAAATtTGGAGTCATCTATTATGACATAAACCAA 189

QY 374 IleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLys 393
Db 190 ATTATTACAACAGCCGTGATGACCCACACTAAACACTTCGATGCTTATGGCCGGTGAAG 249

QY 394 GluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsn 413
Db 250 GAGTCCAATATGAATATTTCGATCTCTTATGTtTCGGATGATGGTGcAGTATGACAAC 309

QY 414 MetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyr 433
Db 310 ATGGGGCGTGTGGTGGCCAAAGAGAGCTAAAGGTGGGACCGCTATGCCAACACGACACGCTAT 369

QY 434 SerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeu 453
Db 370 ACATATGAATATGATGCCGATGGACAACtTTCAGGTAGTCTCCATCAATGACAAGCCCTG 429

QY 454 TrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAla 473
Db 430 TGGAGGTACAGCTATGACCTAAATGGAAACCTGCACCTTCTCAGCCCTGGAAATAGTGCA 489

QY 474 ArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGln 493
Db 490 CGCCTCACCCCGCTACGTATGATAACAGGGACCGCATTACTCGCTTGGCGGATGTACAG 549

QY 494 TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn 513
Db 550 TACAGGATGGACGAGGACGGTtTtTCTCAAGCAGCGGAGGAATGACTACTCTCAATACMAC 609

QY 514 SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArg 533

Mon Aug 16 09:01:11 2004

Db 610 TCAGCAGGCCCTTCTAGTTAGGGTGTACAACAAAAGTGAGCGGTTGGAGCATAACAATATCGG 669

QY 534 TyrAspGlyLeuGlyArgArgVal 541

Db 670 TATGATGGCCTGGGAAGGAGGGTG 693

Search completed: August 14, 2004, 18:04:37
Job time : 6395.74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 917.682 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1760_2300
Perfect score: 2845
Sequence: 1 YYIGADGSLRLLLANGMEVA.....YNRAGSWSVRYDGLGRRV 541

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues 6747726

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1.1868@runat_06082004_112215_29265 -NCPV=6 -ICPV=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_29Jan04:	*
1: geneseqn1980s:	*
2: geneseqn1990s:	*
3: geneseqn2000s:	*
4: geneseqn2001as:	*
5: geneseqn2001bs:	*
6: geneseqn2002s:	*
7: geneseqn2003as:	*
8: geneseqn2003bs:	*
9: geneseqn2003cs:	*
10: geneseqn2004s:	*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2845	100.0	8354	6	ABS52100	Abs52100 Human TEN
2	2845	100.0	8645	6	ABS78652	Abs78652 Human CDN
3	2821.5	99.2	8438	6	ABN85378	Abn85378 Human NOV
4	2233	78.5	3614	4	AAH14183	Aah14183 Human CDN
5	2233	78.5	8473	6	ABQ82345	Abq82345 Human NOV
6	2233	78.5	8487	6	ABQ82346	Abq82346 Human NOV
7	2233	78.5	8645	6	ABQ82344	Abq82344 Human NOV
8	2233	78.5	8675	6	ABQ82343	Abq82343 Human NOV

9	2159	75.9	3270	4	AAH14671	Aah14671 Human CDN
10	2136.5	75.1	9058	7	ACC72051	Acc72051 BCU0205A
11	2136.5	75.1	9695	7	ACC72052	Acc72052 BCU0205B
12	2136.5	75.1	9729	5	AAS14089	Aas14089 Human FCT
13	2136.5	75.1	9729	9	ADB32028	Adb32028 Human FCT
14	2136.5	75.1	9826	5	AAS14085	Aas14085 Human FCT
15	2136.5	75.1	9826	9	ADB32023	Adb32023 Human FCT
16	2052.5	72.1	13202	4	AAK51828	Aak51828 Human pol
17	2048.5	72.0	12879	6	ABK92230	Abk92230 Prostate
18	1744.5	61.3	3312	4	AAK52812	Aak52812 Human pol
19	1254.5	44.1	1227	5	AAS68861	Aas68861 DNA encod
20	999	35.1	812	4	AAH06979	Aah06979 Human CDN
21	982.5	34.5	806	4	AAI23889	Aai23889 Probe #13
22	982.5	34.5	806	4	ABA69008	Aba69008 Human foe
23	982.5	34.5	806	4	AAI49198	Aai49198 Probe #17
24	982.5	34.5	806	4	ABA51018	Aba51018 Human bre
25	982.5	34.5	806	4	ABA35953	Aba35953 Probe #14
26	982.5	34.5	806	4	AAK43120	Aak43120 Human bon
27	982.5	34.5	806	4	AAK17327	Aak17327 Human bra
28	982.5	34.5	806	4	ABS42754	Abs42754 Human liv
29	982.5	34.5	806	5	AAI09491	Aai09491 Probe #94
30	982.5	34.5	806	6	ABS17195	Abs17195 Human gen
31	982.5	34.5	1973	4	AAI14695	Aai14695 Probe #46
32	982.5	34.5	1973	4	ABA56427	Aba56427 Human foe
33	982.5	34.5	1973	4	AAI36063	Aai36063 Probe #47
34	982.5	34.5	1973	4	ABA45900	Aba45900 Human bre
35	982.5	34.5	1973	4	ABA26060	Aba26060 Probe #45
36	982.5	34.5	1973	4	AAK30099	Aak30099 Human bon
37	982.5	34.5	1973	4	AAK04592	Aak04592 Human bra
38	982.5	34.5	1973	4	ABS29749	Abs29749 Human liv
39	982.5	34.5	1973	5	AAI04501	Aai04501 Probe #44
40	982.5	34.5	1973	6	ABS04679	Abs04679 Human gen
41	882.5	31.0	640	4	AAI26459	Aai26459 Human bre
42	764	26.9	843	4	AAH05415	Aah05415 Human CDN
43	764	26.9	2157	4	AAH14096	Aah14096 Human CDN
44	607.5	21.4	487	4	AAI16510	Aai16510 Human bre
45	585.5	20.6	410	5	AAS68859	Aas68859 DNA encod

ALIGNMENTS

RESULT 1

ABS52100

ID ABS52100 standard; DNA; 8354 BP.

XX

AC ABS52100;

XX

DT 05-NOV-2002 (first entry)

XX

DE Human TEN-M4-like gene.

XX

KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBF1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type Ia membrane sushi-containing domain; butyrophilin;
KW single nucleotide polymorphism.

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT variation replace(117,G)

FT /*tag= a

QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 6452 ACCACACCAAGCATTTTGATGCATATGGCAGGATGAAGGAGTGCGAGTATGAGATCTTC 6511
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420
Db 6512 CGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAAG 6571
QY 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6572 GAGCTGAAGTAGGACCTACGCCAATACCACTCGTACTCCTATGAGTATGATGCTGAC 6631
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460
Db 6632 GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC 6691
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6692 AATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTAT 6751
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6752 GACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAAGATGGATGAGGATGSC 6811
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 6812 TTCCTGAGGCAGCGGGCGGTGATATCTTTGAGTACAACCTCAGTGGCCTGCTCATCAAG 6871
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 6872 GCCTACAACCGGGCTGGCAGTGGAGTGTCAAGTACCCTACGATGGCCTGGGGCGGCGC 6931
QY 541 Val 541
Db 6932 GTG 6934
RESULT 2
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX AC ABS78652;
XX AC ABS78652;
DT 16-DEC-2002 (first entry)
XX DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX OS Homo sapiens.
XX OS
XX PN WO200272830-A2.
XX PD 19-SEP-2002.
XX PF 08-FEB-2002; 2002WO-US003715.
XX PR 09-FEB-2001; 2001US-0268111P.
XX PR 23-FEB-2001; 2001US-0271175P.
XX PR 08-MAR-2001; 2001US-0274503P.
XX PR 09-MAR-2001; 2001US-0274552P.
XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX DR WPI; 2002-723356/78.
DR P-PSDB; ABG97359.
XX New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX Claim 5; Page 175-178; 181pp; English.
XX The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides and a
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;
Alignment Scores:
Pred. No.: 2.93e-278 Length: 8645
Score: 2845.00 Matches: 541
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1760_2300 (1-541) x ABS78652 (1-8645)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db 5361 TACTACATCGGGGCGGATGGCTCTTGGGCTGTGTGGCCAAACGGCATGGAGGTGGCG 5420
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5421 CTGCAGACTGAGCCCCACTTGTGGCTGGCACCGTCAACCCACCGTGGCAAGAGGAAT 5480
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5481 GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGACGAG 5540
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5541 GCTCGGGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGGTGCAACCCGAAATCTCCTA 5600

QY	81	SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe	100	Db	6681	GGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGACCTC	6740
Db	5601	TCCTGGACTTTGATCGCGTAACACGCACAGAGAAGATCTATGATGACCACCGCAAGTTC	5660	QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
QY	101	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg	120	Db	6741	AATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACCTACGGTAT	6800
Db	5661	ACCCCTCGGATCTGTACAGCACGCGGGCGCGCCAGCCCTCTGGTCACCCAGCAGCAGG	5720	QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrIlysMetAspGluAspGly	500
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly	140	Db	6801	GACATCCGCGACCGCATCACTCGGTGGTGACGTGCAATACAAAGATGGATGAGGATGGC	6860
Db	5721	CTGAATGGTCTCAACGTGACATACTCCCTCGGGGTTACATTGCTGGCATCCAGAGGGGC	5780	QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
QY	141	IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla	160	Db	6861	TTCTGTAGCGAGCGGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAG	6920
Db	5781	ATCATGTCTGAAGAATGGAATACGACCGCGCGCGCATCACATCCAGGATCTTCGCT	5840	QY	521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
QY	161	AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer	180	Db	6921	GCCTACAACCGGCTGGCAGCTGGAGTGTGAGTACCGTACCGTACGATGGCGGGCGCGC	6980
Db	5841	GATGGGAAGACATGGAGCTACACATCTTAGAGAAGTCCATGGTGTGCTACTACACAGC	5900	QY	541	Val 541	
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro	200	Db	6981	GTG 6983	
Db	5901	CAGAGGCAGTATATCTTTGAGTTCGACAAGAATGACCGCCTCTCTTCTGTGACGATGCC	5960	RESULT 3			
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr	220	ID	ABN85378		
Db	5961	AACGTGGCGCGGACACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTAT	6020	XX	ABN85378 standard; DNA; 8438 BP.		
QY	221	GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu	240	AC	ABN85378;		
Db	6021	CAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACTTCACTGAGGATGGGACCTCCTT	6080	XX	21-OCT-2002 (first entry)		
QY	241	HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys	260	DE	Human NOV1, TEN-M4 like protein, coding sequence.		
Db	6081	CACACCTTCTACCTGGGCACTGGCCGAGGGTGATATACAAAGTATGGCAAACCTGCAAAAG	6140	XX	Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;		
QY	261	LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly	280	KW	Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;		
Db	6141	CTGGCAGAGACGCTCTATGACACACCACCAAGGTTCAGTTTCACTATGACGAGACGGCAGGC	6200	KW	Antiaesthetic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;		
QY	281	MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln	300	KW	Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;		
Db	6201	ATGCTGAAGACCATCAACCTACAGAATGAGGGCTTCACTGACCATCCGCTACCGTCTCAG	6260	KW	Gene Therapy; NOV; cancer; heart disease; inflammation;		
QY	301	IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla	320	KW	autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;		
Db	6261	ATTGGGCCCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACGCC	6320	KW	asthma; IgA nephropathy; cirrhosis; arthritis; Alzheimer's disease;		
QY	321	ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn	340	KW	infection; stroke; muscular dystrophy; epilepsy; wasting disorder;		
Db	6321	CGTTTGTACTACAACATATGACACACAGCTTCCGGGTGACCAAGCATGACGGCTGTGATCAAC	6380	XX	TEN-M4 like protein; chromosome 11; gene; ds.		
QY	341	GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln	360	OS	Homo sapiens.		
Db	6381	GAGACCCCACTGCCATTGATCTCTATCGCTATGATGATGTGTGAGCAAGACAGAGCAG	6440	XX	Location/Qualifiers		
QY	361	PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet	380	Key	4. .8395		
Db	6441	TTTGGGAAGTTGGTGTCTATTACTATGACATTAAACCAGATCATCACACAGCTGTCTATG	6500	CDS	/*tag= a		
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400	FT	/trans_except= (pos: 1138. .1147,aa:Met)		
Db	6501	ACCCACACCAAGCATTTTGTGATGATATGGCAGGATGAAGGAAGTGCAGTATGAGATCTTC	6560	FT	/product= "NOV1 protein"		
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420	XX	WO200255704-A2.		
Db	6561	CGCTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGGCGAGTAGTGAAGAAG	6620	XX	18-JUL-2002.		
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440	XX	09-JAN-2002; 2002WO-US0000554.		
Db	6621	GAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGAGTATGATGCTGAC	6680	XX	09-JAN-2001; 2001US-0260417P.		
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu	460	XX	10-JAN-2001; 2001US-0260831P.		
				XX	28-FEB-2001; 2001US-0272338P.		
				XX	09-MAR-2001; 2001US-0274876P.		
				XX	18-APR-2001; 2001US-0284704P.		
				PA	(CURA-) CURAGEN CORP.		
				XX	Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;		
				PI	Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;		
				PI	Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;		
				PI	Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;		
				PI	Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;		
				PI	Gunther E, Stone DJ;		
				XX	WPI; 2002-590674/63.		
				DR			

DR P-PSDB; ABB98401.

XX NOVX polypeptides and encoding polynucleotides, useful for preventing or

PT treating NOVX-associated disorders e.g. cancer, inflammation, or

PT Alzheimer's disease, and in chromosome mapping, tissue typing or

PT pharmacogenomics.

XX

PS Claim 9; Page 8-9; 358pp; English.

XX

CC The present sequence is a coding sequence for a NOV protein. The NOV

CC proteins and coding sequences are useful for treating or preventing NOV-

CC associated disorders or in the manufacture of a medicament for treating

CC the disorders, such as cancer, heart disease, inflammation, autoimmune

CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,

CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections

CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,

CC and other wasting disorders associated with chronic diseases. NOV1 is a

CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11

XX

SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.95e-276 Length: 8438

Score: 2821.50 Matches: 541

Percent Similarity: 98.72% Conservative: 0

Best Local Similarity: 98.72% Mismatches: 0

Query Match: 99.17% Indels: 7

DB: 6 Gaps: 2

US-10-029-020-14_COPY_1760_2300 (1-541) x ABN85378 (1-8438)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db 5318 TACTACATCGGGCCGATGGCTCTTGGGCTGCTGGCTGGCCAAACGGCATGGAGTGGCG 5377

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 5378 CTGCAGACTGAGCCCACTTGCTGGCTGGCACCGTCAACCCACCGTGGGCAAGAGGAAT 5437

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60

Db 5438 GTCACGCTGCCCATCGACAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAG 5497

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArg-----ValHisAsnArg 77

Db 5498 GCTCGGGCCAGGTCACTGTCTTTGGGCGCCGGTGTCTCCAGGTTCAACAACCGA 5557

QY 78 AsnLeuLeuSerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHis 97

Db 5558 AATCTCCTATCTCTGGACTTTTGATCGCGTAAACACGCACAGAGAGATCTATGATGACCAC 5617

QY 98 ArgLysPheThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerPro 117

Db 5618 CGCAAGTTACACCTTCGGATTCTGTACGACCAGGGGGGGCGCCAGCCTCTGGTCACCC 5677

QY 118 SerSerArgLeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIle 137

Db 5678 AGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCTGGGGGTACATTGCTGGCATC 5737

QY 138 GlnArgGlyIleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArg 157

Db 5738 CAGAGGGGCATCATGCTCTGAAAGATGGAATACGACCAGCGCGCGCATCATCCAGG 5797

QY 158 IlePheAlaAspGlyLysThrTrpSerTyrThrTyrLeuGluLys-----Ser 173

Db 5798 ATCTTCGTGATGGGAAGACATGGAGCTACACATACTTAGAAGGCGAGGTGTCCAGTCC 5857

QY 174 MetValLeuLeuHisSerGlnArgGlnTyrIlePheGluPheAspLysAsnAspArg 193

Db 5858 ATGGTGTGCTACTACACAGCCAGGCGAGTATATCTTTAGTTCGACAAGATGACCGC 5917

QY 194 LeuSerSerValThrMetProAsnValAlaArgGlnThrLeuGluThrIleArgSerVal 213

Db 5918 CTCTCTTCTGTGACGATGCCCAACGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTG 5977

QY 214 GlyTyrTyrArgAsnIleTyrGlnProProGluGlyAsnAlaSerValIleGlnAspPhe 233

Db 5978 GGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCATACAGGACTTC 6037

QY 234 ThrGluAspGlyHisLeuLeuHisThrPheTyrLeuGlyThrGlyArgArgValIleTyr 253

Db 6038 ACTGAGGATGGGCACCTCCCTTCACACCTTCTACCTGGGCACCTGGCCGAGGGTGATATAC 6097

QY 254 LysTyrGlyLysLeuSerLysLeuAlaGluThrLeuTyrAspThrThrLysValSerPhe 273

Db 6098 AAGTATGGCAAACCTGTCAAAGCTGGCAGAGACGCTCTATGACACCACCAAGGTCAGTTTC 6157

QY 274 ThrTyrAspGluThrAlaGlyMetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThr 293

Db 6158 ACCTATGACGAGACGGCAGGCATGCTGAAGACCATCACTACAGAATGAGGGCTTCACC 6217

QY 294 CysThrIleArgTyrArgGlnIleGlyProLeuIleAspArgGlnIlePheArgPheThr 313

Db 6218 TGCACCATCCGCTACCGTCAGATTGGGCCCTTGATTGACCGACAGATCTTCCGCTTCACT 6277

QY 314 GluGluGlyMetValAsnAlaArgPheAspTyrAsnTyrAspAsnSerPheArgValThr 333

Db 6278 GAGGAAGGCATGGTCAACGCCCGTTTGTGACTACAACTATGACAACAGCTTCCGGGTGACC 6337

QY 334 SerMetGlnAlaValIleAsnGluThrProLeuProIleAspLeuTyrArgTyrAspAsp 353

Db 6338 AGCATGCAGGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTATCGCTATGATGAT 6397

QY 354 ValSerGlyLysThrGluGlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGln 373

Db 6398 GTGTGAGGCAAGACAGACAGCAGTTTGGGAAGTTTGGTGTCACTTACTATGACATTAACCA 6457

QY 374 IleIleThrThrAlaValMetThrHisThrLysHisPheAspAlaTyrGlyArgMetLys 393

Db 6458 ATCATCACCAAGTGTATGATGACCCACACCAAGCATTTTGTGATGATATGCGAGGATGAAG 6517

QY 394 GluValGlnTyrGluIlePheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsn 413

Db 6518 GAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACCCGTCAGTATGATAAC 6577

QY 414 MetGlyArgValValLysLysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyr 433

Db 6578 ATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTAC 6637

QY 434 SerTyrGluTyrAspAlaAspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeu 453

Db 6638 TCCTATGATATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAGCCACTC 6697

QY 454 TrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAla 473

Db 6698 TGGCGCTACAGTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTGCA 6757

QY 474 ArgLeuThrProLeuArgTyrAspIleArgAspArgIleThrArgLeuGlyAspValGln 493

Db 6758 CGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAA 6817

QY 494 TyrLysMetAspGluAspGlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsn 513

Db 6818 TACAAGATGGATGAGGATGGCTTCTGTAGGCAGCGGGCGGTGATATCTTTGAGTACAAC 6877

QY 514 SerAlaGlyLeuLeuIleLysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArg 533

Db 6878 TCAGTGGCCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGTCAAGTACCGC 6937

QY 534 TyrAspGlyLeuGlyArgArgVal 541

Db 6938 TACGATGGCCTGGGGCGGCGCGTG 6961

RESULT 4

AAH14183

ID AAH14183 standard; cDNA; 3614 BP.

XX AC AAH14183;
XX XX
DT 26-JUN-2001 (first entry)
XX XX
DE Human cDNA sequence SEQ ID NO:11430.
XX XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX XX
OS Homo sapiens.
XX XX
PN EP1074617-A2.
XX XX
PD 07-FEB-2001.
XX XX
PF 28-JUL-2000; 2000EP-00116126.
XX XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX XX
PA (HELI-) HELIX RES INST.
XX XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX XX
DR WPI; 2001-318749/34.
XX XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX XX
PS Claim 8; SEQ ID NO 11430; 2537pp + Sequence Listing; English.
XX XX
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the nucleotide sequence comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
XX XX
SQ Sequence 3614 BP; 967 A; 898 C; 928 G; 821 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.56e-216 Length: 3614
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AAH14183 (1-3614)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db 272 TACCAGATTGGTTATGACGGCTCCCTCAGAATTATCTACGCCAGTGGCTTGACTCACAC 331
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 332 TACCAAAACAGAGCCGCGACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAAGAAAC 391
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 392 ATGACTTTGGCTGGCGAGAACCGTCAAAACTTGGTGAATGGAGATTCCGAAAAGAGCAA 451
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 452 GCCCAAGGGAAAGTCAATGTCCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTT 511
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 512 TCAGTTGACTTTGATCGAACCAACAAAGACAGACAAAAGATCTATGACGACCCCGTAATTT 571
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 572 CTACTGAGGATCGCTACGACACGTCCTGGGACCCGACTCTCTGGCTGCCAAGCAGCAAG 631
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 632 CTGATGGCCGTCATGTCACTATTCAATCCACAGGTCAAATTGCCAGCATCCAGCGAGGC 691
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 692 ACCACTAGCGAGAAAGTAGATTATGACGGGACAGGGGAGGATCGTGTCTCGGCTTTGCT 751
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 752 GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 811
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 812 CAGCGGCAGTACATCTCGAATACGATATGTGGACCGCCTGTCTGCCATCACCATGCC 871
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 872 AGTGTGGCTCGCCACACCATGACAGCATCCGATCCATGGCTACTACCGCAACATATAC 931
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 932 AACCCCGGAAAGCAACGCTCCATCATCACGGACTACAGAGGAGGAGGCTGCTTCTA 991
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 992 CAAACAGCTTTCTTGGGTACAGTCGGAGGGTCTTATTCAAATACAGAAAGGAGACTAGG 1051
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 1052 CTCACAGAAATTTTATATATGATAGCACAGAGTCAGTTTACCTATGATGAACAGCAGGA 1111
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 1112 GTCCTAAAGACAGTAAACCTCCAGAGTGATGGTTTTATTTCACCATTAGATACAGGCAA 1171
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 1172 ATTGGTCCCCTGATTGACAGGCGAGATTTTCCGCTTTTAGTGAAGATGGGATGTAATGCA 1231
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 1232 AGATTTGACTATAGCTATGACAAACAGCTTTCGAGTGACCCAGCATGCGGGTGTGATCAAT 1291
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 1292 GAAACGCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGCAAAAGTTGAGCAG 1351
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrAlaValMet 380

Db	1352	TTTGAAAGTTTGGAGTTATATATTATGATATTAAACGATCAATTTACAGCTGTAATG	1411
QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
Db	1412	ACCTATACGAAGCACTTTTGATGCTCATGGCGGTATCAAGGAGATTCAATATGAGATATTC	1471
QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallLysLys	420
Db	1472	AGGTCGCTCATGTACTGGATTACAATTCACTATGATAACATGGGTGGGTAAACCAAGAGA	1531
QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
Db	1532	GAGATTAAAAATAGGGCCCTTTTGCCCAACACCACCAAAATATGCTTATGAATATGATGTTGAT	1591
QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu	460
Db	1592	GGACAGCTCCAAACAGTTTACTCAATGAAAGATAATGTGGCGGTACAACACTACGATCTG	1651
QY	461	AsnGlyAsnLeuHisLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
Db	1652	AATGGAACCTCCATTTACTGAACCCCAAGTAACAGTGGCGCTGTGACACCCCTTCGCTAT	1711
QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
Db	1712	GACCTGCGAGACAGAATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGATGGT	1771
QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
Db	1772	TTCTTACGTCAAAGGGGCACGGAATCTTTGAATATAGCTCCAAGGGCTTCTAACTCGA	1831
QY	521	AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
Db	1832	GTTTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCGTTATGACGGCCTGGGAAGGCGT	1891
QY	541	Val 541	
Db	1892	GTT 1894	
RESULT 5			
ABQ82345			
ID	ABQ82345	standard; cDNA; 8473 BP.	
XX			
AC	ABQ82345;		
XX			
DT	17-DEC-2002	(first entry)	
XX			
DE	Human NOV15c	encoding cDNA SEQ ID NO:39.	
XX			
KW	Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disease; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	258..8144	
FT		/*tag= a	
FT		/product= "NOV15c"	
XX			
PN	WO200262999-A2.		
XX			
PD	15-AUG-2002.		
XX			
PF	31-DEC-2001; 2001WO-US049976.		
XX			

PR	29-DEC-2000;	2000US-0258928P.
PR	02-JAN-2001;	2001US-0259415P.
PR	04-JAN-2001;	2001US-0259785P.
PR	20-FEB-2001;	2001US-0269814P.
PR	09-MAR-2001;	2001US-0279863P.
PR	29-MAR-2001;	2001US-0279832P.
PR	29-MAR-2001;	2001US-0279833P.
PR	13-APR-2001;	2001US-0283889P.
PR	18-APR-2001;	2001US-0284447P.
PR	25-APR-2001;	2001US-0286683P.
PR	29-MAY-2001;	2001US-0294080P.
PR	16-AUG-2001;	2001US-0312915P.
PR	17-AUG-2001;	2001US-0313325P.
PR	17-SEP-2001;	2001US-0322699P.
PR	26-NOV-2001;	2001US-0333350P.
XX		
PA	(CURA-) CURAGEN CORP.	
XX		
PI	SpYTEK KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;	
PI	Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;	
PI	Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;	
PI	Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;	
PI	Gunther E, Smithson G, Millet I, Macdougall JR;	
XX		
DR	WPI; 2002-732706/79.	
DR	P-PSDB; ABP53588.	
XX		
PT	New NOVX polypeptides and polynucleotides useful for treating NOVX-	
PT	associated disorders, such as cancers, neurological disorders, disorders	
PT	of vesicular transport, gastrointestinal disorders, and autoimmune	
PT	diseases.	
XX		
PS	Claim 8; Page 119-121; 444pp; English.	
XX		
CC	The present invention describes novel human proteins designated NOVX,	
CC	where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,	
CC	cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,	
CC	antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,	
CC	immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,	
CC	protozoacide and antihelminthic activities, and can be used in gene	
CC	therapy. The NOVX proteins, nucleotides or antibodies can be used in the	
CC	manufacture of a medicament for treating a syndrome associated with a	
CC	human disease selected from NOVX-associated disorder, such as cancers	
CC	(e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,	
CC	ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,	
CC	ischaemic cerebrovascular disease, Alzheimer's disease or pick's	
CC	disease), disorders of vesicular transport (e.g. cystic fibrosis,	
CC	diabetes mellitus, Grave's disease, or goitre), gastrointestinal	
CC	disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),	
CC	autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic	
CC	anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic	
CC	and protozoal infections. The NOVX proteins can be used as immunogens to	
CC	produce antibodies and as vaccines. The NOVX nucleotide sequences may be	
CC	used in chromosome mapping, identifying individuals from minute	
CC	biological samples (tissue typing), and in forensic identification of a	
CC	biological sample. The present sequence encodes human NOV15c, which is	
CC	located on chromosome 4	
XX		
SQ	Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;	

Alignment Scores:

Pred. No.:	4.98e-216	Length:	8473
Score:	2233.00	Matches:	404
Percent Similarity:	88.91%	Conservative:	77
Best Local Similarity:	74.68%	Mismatches:	60
Query Match:	78.49%	Indels:	0
DB:	6	Gaps:	0

US-10-029-020-14_COPY_1760_2300 (1-541) x ABQ82345 (1-8473)

QY	1	TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla	20
Db	5130	TACCAGATTGGTTATACGGCTCCCTCAGAAATTATCTACGCCAGTGGCCTGGACTCACAC	5189

QY	21	LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn	40	QY	381	ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe	400
Db	5190	TACCAACAGAGCCGACGCTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAGAAAC	5249	Db	6270	ACCTATACGAAGCACTTTTGATGCTCATGGCCGTATCAAGAGATTCAATATGAGATATTC	6329
QY	41	ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln	60	QY	401	ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys	420
Db	5250	ATGACTTTGCCCTGGCGAAGACGGTCAAACTTGGTGAATGGAGATTCCGAAAGAGCAA	5309	Db	6330	AGGTCGCTCATGTACTGGATTACAATTCAATCAGTATGATAACATGGTCCGGTAACCAAGAGA	6389
QY	61	AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu	80	QY	421	GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp	440
Db	5310	GCCCAAGGGAAGTCAATGTCTTTGGCCGAAGCTCAGGGTTAATGGCAGAAACCTCCTT	5369	Db	6390	GAGATTAATAATAGGCCCTTTGCCAACACACCACCAATAATGCTTATGARTATGTTGAT	6449
QY	81	SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe	100	QY	441	GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu	460
Db	5370	TCAGTTGACTTTGATCGAACAAACAAAGACAGAAAAGATCTATGACGACACCCGTAAATT	5429	Db	6450	GGACAGCTCCAAACAGTTTACCTCAATGARAAGATAATGTGGCGGTACAACTACGATCTG	6509
QY	101	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPsrProSerSerArg	120	QY	461	AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr	480
Db	5430	CTACTAGGATCGCCTACGACACGCTCTGGGCAACCGACTCTCTGGTCTGCCAAGCAGCAAG	5489	Db	6510	AATGGAAACCTCCATTACTGAACCCCAAGTAACAGTGCCTCTGACACCCCTTCGCTAT	6569
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly	140	QY	481	AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly	500
Db	5490	CTGATGGCCGCTCAATGTCACTTATTCATCCACAGTCAAAATGCCAGCATCCAGCGAGGC	5549	Db	6570	GACCTGCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGTTGGATGAAGATGGT	6629
QY	141	IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla	160	QY	501	PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys	520
Db	5550	ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGCTTTTGCT	5609	Db	6630	TTCTTACGTCAAGGGGCACGGAAATCTTTGAATATAGTCCCAAGGGGCTTCTAACTCGA	6689
QY	161	AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer	180	QY	521	AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg	540
Db	5610	GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGTTCTTCTGCTTCATAGC	5669	Db	6690	GTTTACAGTAAAGGCAGTGGCTGGACAGTGTACTACCGTTATGACGGCCTGGGAAGGCGT	6749
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro	200	QY	541	Val 541	
Db	5670	CAGCGCAGTACATCTTCGAATACGATATGTGGGACCGCTGTCTGCCATCACCATGCC	5729	Db	6750	GTT 6752	
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr	220	RESULT 6			
Db	5730	AGTGTGGCTCGCCACACCATGACAGACCATCCGATCCATGGCTACTACCGCAACATATAC	5789	ABQ82346			
QY	221	GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu	240	ID	ABQ82346	standard; cDNA; 8487 BP.	
Db	5790	AACCCCGGAAAGCAACGCCTCCATCATCAGGACTACAGGACTACAAAGGAGGGTGCTTCTA	5849	XX	ABQ82346;		
QY	241	HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys	260	XX	17-DEC-2002	(first entry)	
Db	5850	CAAAACAGCTTTCTGGGTACAAGTCGAGGGTCTTATTCAAATACAGAAGGCAGACTAGG	5909	DE	Human NOV15d	encoding cDNA SEQ ID NO:41.	
QY	261	LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly	280	XX	Human; NOVX;	cytostatic; neuroprotective; anticonvulsant; cardiovascular;	
Db	5910	CTCTCAGAAATTTTATATGATAGCACAAAGAGTCAGTTTTTACCTATGATGAACAGCAGGA	5969	KW	cerebroprotective;	nootropic; antidiabetic; antiinflammatory; fungicide;	
QY	281	MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln	300	KW	antirheumatic;	antiarthritic; immunosuppressive; antiallergic; virucide;	
Db	5970	GTCCTAAAGACAGTAAACCTCCAGAGTGTGTTTTTATTTGCAACCATGATACAGGCAA	6029	KW	antianaemic;	antibacterial; protozoacide; antihelminthic; gene therapy;	
QY	301	IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla	320	KW	cancer;	leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;	
Db	6030	ATTGTCCTCCCTGATTGACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGTTAAATGCA	6089	KW	stroke;	ischemic cerebrovascular disease; Alzheimer's disease; allergy;	
QY	321	ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn	340	KW	pick's disease;	vesicular transport disease; cystic fibrosis; goitre;	
Db	6090	AGATTTGACTATAGCTATGATCTGTATCAGTTTGATGACATTTCTGGCAAAAGTTGAGCAG	6149	KW	diabetes mellitus;	Grave's disease; gastrointestinal disorder; infection;	
QY	341	GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln	360	KW	ulcerative colitis;	gastric disorder; duodenal disorder; autoimmune	
Db	6150	GAAACGCCACTGCTTATGATCTGTATCAGTTTGATGACATTTCTGGCAAAAGTTGAGCAG	6209	KW	rheumatoid arthritis;	gene; chromosome 4; ss.	
QY	361	PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet	380	XX	Homo sapiens.		
Db	6210	TTTGGAAAGTTTGGAGTTATATATATGATATTAACCAAGATCATTTCTACAGCTGTAATG	6269	OS	Location/Qualifiers		
				XX	Key	299. .8140	
				XX	CDS	/tag= a	
				FT		/product= "NOV15d"	
				XX			
				PN	WO200262999-A2.		
				XX			
				PD	15-AUG-2002.		
				XX			
				PF	31-DEC-2001; 2001WO-US049976.		
				XX			
				PR	29-DEC-2000; 2000US-0258928P.		
				PR	02-JAN-2001; 2001US-0259415P.		

PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53589.
XX
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 123-125; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytosstatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15d, which is
CC located on chromosome 4
XX
SQ Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;

Alignment Scores:
Pred. No.: 4.99e-216 Length: 8487
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x ABQ82346 (1-8487)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValala 20
Db 5126 TACCAGATTGGTTATGACGGCTCCCTCAGAATTATCTACGCCAGTGGCTGGACTCACAC 5185
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 5186 TACCAAAACAGAGCGCGACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAGAAAC 5245
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5246 ATGACTTTGCCTGGCGAGAACGGTCAAAACTTGGTGAATGGAGATTCCGAAAAGAGCAA 5305
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5306 GCCCAAGGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGTTAATGGCAGAAACCTCCTT 5365
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspHisArgLysPhe 100
Db 5366 TCAGTTGACTTTGATCGAAACAACAAGACAGAAAAGATCTATGACGACCACCGTAAATTT 5425
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 5426 CTACTGAGGATCGCTACGACACAGTCTGGGCACCCGACTCTCTGGCTGCCACAGCAAG 5485
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5486 CTGATGGCCGTCATGTACCTATTTCATCCACAGTCAAATTGCCAGCATCCACGGAGGC 5545
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5546 ACNACTAGCGAGAAAGTAGATTATGACGGACAGGGAGGATCGTGTCTCGGCTCTTGCT 5605
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5606 GATGGTAAACATGGAGTTACACATATTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 5665
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5666 CAGCGGCAGTACATCTTCGAATACGATATGTGGGACCGCCTGTCTGCCATCACCATGCCC 5725
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5726 AGTGTGGTCCGCCACACCATGCAGACCATCCGATCCATTGGCTACTACCGCAACATATAC 5785
QY 221 GlnProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 5786 AACCCCCCGAAAGCAACGCCTCCATCATCAGGACTACAAACGAGGAAGGGTGCTTCTA 5845
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 5846 CAAACAGCTTCTTGGGTACAAGTCGAGGGTCTTATTCAAATACAGAAGGCAGACTAGG 5905
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 5906 CTCTCAGAAATTTTATATGATGACACAAGAGTCAGTTTTACCTATGATGAAACAGCAGGA 5965
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 5966 GTCCTAAAGACAGTAAACCTCCAGAGTGATGGTTTTATTTCACCATTAGATACAGGCAA 6025
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6026 ATTGTCCTCCCTGATTGACAGGCAGATTTTCCGCTTTAGTGAAGATGGGATGGTAAATGCA 6085
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6086 AGATTTGACTATAGCTATGACAACAGCTTTCGAGTGACACCATGCAGGGTGTGATCAAT 6145
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 6146 GAAACGCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG 6205
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 6206 TTTGGAAAAGTTTGGAGTTATATATATGATATTAAACCAGATCATTTCTACAGCTGTAATG 6265
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400

Db 6266 ACCTATACGAAGCACCTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC 6325

Qy 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallLysLys 420

Db 6326 AGGTCGCTCATGTACTGGATTACAATTCAGTATGATGATGGGTGGGTAAACCAAGAGA 6385

Qy 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaasp 440

Db 6386 GAGATTAAAAATAGGGCCCTTTGCCAACACACCACCAATATGCTTATGAATATGATTTGAT 6445

Qy 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeu 460

Db 6446 GGACAGCTCCAAACAGTTTACCTCAATGAAAAAGATAATGTGGCGGTACAACACTACGATCTG 6505

Qy 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480

Db 6506 AATGGAACCTCCATTACTGAACCCCAAGTAACAGTGGCGGTCTGACACCCCTTCGCTAT 6565

Qy 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500

Db 6566 GACCTGCAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGATGGT 6625

Qy 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuLys 520

Db 6626 TTCCTACGTCAAAGGGGCACGGAAATCTTTGAATATAGCTCCAAGGGGTCTTAACCTGA 6685

Qy 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540

Db 6686 GTTTACAGTAAAGGCAGTGGCTGGACAGTGATCTACCGTTATGACGGCGCTGGGAAGCGT 6745

Qy 541 Val 541

Db 6746 GTT 6748

RESULT 7

ABQ82344

ID ABQ82344 standard; cDNA; 8645 BP.

AC ABQ82344;

XX 17-DEC-2002 (first entry)

DE Human NOV15b encoding cDNA SEQ ID NO:37.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; Pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

OS

XX

PH Key Location/Qualifiers

CDS 151..8316

FT /*tag= a

FT /product= "NOV15b"

FT

XX

WO200262999-A2.

15-AUG-2002.

31-DEC-2001; 2001WO-US049976.

29-DEC-2000; 2000US-0258928P.

02-JAN-2001; 2001US-0259415P.

04-JAN-2001; 2001US-0259785P.

20-FEB-2001; 2001US-0269814P.

PR 09-MAR-2001; 2001US-0279863P.

PR 29-MAR-2001; 2001US-0279832P.

PR 29-MAR-2001; 2001US-0279833P.

PR 13-APR-2001; 2001US-0283889P.

PR 18-APR-2001; 2001US-0284447P.

PR 25-APR-2001; 2001US-0286683P.

PR 29-MAY-2001; 2001US-0294080P.

PR 16-AUG-2001; 2001US-0312915P.

PR 17-AUG-2001; 2001US-0313325P.

PR 17-SEP-2001; 2001US-0322699P.

PR 26-NOV-2001; 2001US-0333350P.

(CURA-) CURAGEN CORP.

PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

PI Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.

DR P-PSDB; ABP53587.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.

PS Claim 8; Page 114-117; 444pp; English.

XX The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence encodes human NOV15b, which is located on chromosome 4

XX

SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 5.12e-216 Length: 8645

Score: 2233.00 Matches: 404

Percent Similarity: 88.91% Conservative: 77

Best Local Similarity: 74.68% Mismatches: 60

Query Match: 78.49% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x ABQ82344 (1-8645)

Qy 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20

Db 5302 TACCAGATTGGTTATGACGGCTCCCTCAGAAATTATCTACGCCAGTGGCTGCACAC 5361

Qy 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40

Db 5362 TACCAACAGAGCCGACGCTTCTGGCTGGCACCGCTAATCCGCGGTGCCAAAGAAC 5421

Qy 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrPArgGlnArgLysGluGln 60
Db 5422 ATGACTTTGCCTGGCGAAGACGGTCAAAACTTGGTGAATGGAGATTCCGAAAAGAGCAA 5481

Qy 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5482 GCCCAAGGAAAGTCAATGTCTTTGGCCGAAGCTCAGGGTTAATGGCAGAAACCTCCTT 5541

Qy 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5542 TCAGTTGACTTTGATCGAACAACAAGACAGAAAAGATCTATGACGACCACCGTAAATTT 5601

Qy 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrPArgProSerSerArg 120
Db 5602 CTACTGAGGATCGCCTACGACAGCTCTGGGCACCCGACTCTCTGGTGCACAGCAAG 5661

Qy 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5662 CTGATGGCCGTCAATGTCACTATTTCATCCACAGGTCAATTCGCCAGCATCCAGCGAGC 5721

Qy 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5722 ACCACTAGCGAGAAAGTAGATTATGACGGACAGGGGAGGATCGTGTCTCGGGTCTTTGCT 5781

Qy 161 AspGlyLysThrTyrSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5782 GATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 5841

Qy 181 GlnArgGlnTyrIlePheGluPheAspLysAsnArgLeuSerSerValThrMetPro 200
Db 5842 CAGCGGCAGTACATCTCGAATACGATATGTGGACCGCCTGTCTGCCATCACCATGCCC 5901

Qy 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5902 AGTGTGGCTCGCACACCATCGAGACCATCCGATCCATTTGGCTACTACCGCAACATATAC 5961

Qy 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 5962 AACCCCCCGGAAAGCAACGCCTCCATCATCAGGACTACAACAGGAGGGCTGCTTCTA 6021

Qy 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6022 CAAACAGCTTCTTGGGTACAAGTCGGAGGTCTTATTCAAATACAGAAGGCAGACTAGG 6081

Qy 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6082 CTCTCAGAAATTTATATGATAGCACAAAGAGTCAGTTTACCTATGATGAAACAGCAGGA 6141

Qy 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6142 GTCTTAAAGACAGTAAACCTCCAGAGTGATGGTTTATTGTGCACCATAGATACAGGCAA 6201

Qy 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6202 ATTGTCCTGATTGACAGGCAGATTTCGCTTTAGTGAAGATGGGATGGTAAATGCA 6261

Qy 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6262 AGATTGACTATAGCTATGACAACAGCTTTTCGAGTGACCAGCATGCGGGTGTGATCAAT 6321

Qy 341 GluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGluGln 360
Db 6322 GAAACGCCACTGCCTATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAGTTGAGCAG 6381

Qy 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 6382 TTTGGAAAGTTTGGAGTTATATATTATGATATTAAACCAGATCATTTCTACAGCTGTAATG 6441

Qy 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 6442 ACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTC 6501

Qy 401 ArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValVallLysLys 420
Db 6502 AGTTCGCTCATGTACTGGATTACAATTCAGTATGATAACATGCGTCGGGTAAACCAAGAGA 6561

Qy 421 GluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6562 GAGATTAAAATAGGGCCCTTTTGCCAACACCACCACCAATATGCTTATGAATATGATGTTGAT 6621

Qy 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrPArgTyrSerTyrAspLeu 460
Db 6622 GGACAGCTCCAAACAGTTTACTTCAATGAAAAGATAATGTGGCGGTACAACACGATCTG 6681

Qy 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6682 AATGGAACCTCCATTTACTGAACCCCAAGTAACAGTGCCTGATGTTCAATATCGCTTAT 6741

Qy 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6742 GACCTGCGAGACAGATCACTCGACTGGGTGATGTTCAATATCGTTGGATGAAGATGGT 6801

Qy 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 6802 TTCCTACGTCAAAGGGGCACGGAATCTTTGAATATAGTCTCAAGGGGCTTCTAACTCGA 6861

Qy 521 AlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 6862 GTTTACAGTAAAGGCAGTGGCTGGACAGTGTATACCGTTATGACGGCCTGGGAAGGCGT 6921

Qy 541 Val 541
Db 6922 GTT 6924

RESULT 8
ABQ82343
ID ABQ82343 standard; cDNA; 8675 BP.
XX
AC ABQ82343;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15a encoding cDNA SEQ ID NO:35.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; Pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 151..8328
FT /*tag= a
FT /product= "NOV15a"
FT /transl_except= (pos:1249..1251,aa:Ser)
XX
WO200262999-A2.
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.

PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53586.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 110-112; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytotstatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4
XX
SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 5.14e-216 Length: 8675
Score: 2233.00 Matches: 404
Percent Similarity: 88.91% Conservative: 77
Best Local Similarity: 74.68% Mismatches: 60
Query Match: 78.49% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x ABQ82343 (1-8675)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValala 20
Db 5314 TACCAGATTGGTTATGACGGCTCCCTCAGAAATTATCTACGCCAGTGGCTGCACAC 5373
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5374 TACCAAACAGACCGCACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAGAAAC 5433

QY 41 ValThrLeuProIleAspAenGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5434 ATGACTTTTGCTGGCGAGAACGGTCAAAACTTGGTGAATGGAGATTCGGAAGAGCAA 5493
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5494 GCCCAAGGGAAGTCAATGTCCTTTGGCCCAAGCTCAGGGTTAATGGCAGAAAYCTCCTT 5553
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5554 TCAGTTGACTTTTGATCGAACCAAGACAGAAARARATCTATGACGACCCGTAATTT 5613
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 5614 CTACTGAGGATCGCTACGACACGCTCTGGGCACCCGACTCTCTGGCTGCCAAGCAAG 5673
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5674 CTGATGGCCGTCATGTCACCTATTTCATCCACAGGTCAAATTCAGCATCCAGCGAGGC 5733
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5734 ACCACTAGCGAGAAAGTAGATTATGACGACAGGGGAGGATCGTGTCTCGGGTCTTTGCT 5793
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5794 GATGGTAAACATGGAGTTACACATATTAGAAAAGTCCATGGTTCTTCTGCTTCATAGC 5853
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5854 CAGCGGCAGTACATCTTCGAATACGATATGTGGGACCGCTGTCTGCCATCACCATGCCC 5913
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 5914 AGTGTGGCTCGCCACACCATGCGACCATCCGATCCGATCCATGGTACTACCGCAACATATAC 5973
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db 5974 AACCCCCGGGAAGCAACGCTCCATCATCGGACTACCAACGAGGAAGGGTGTCTCTA 6033
QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db 6034 CAAACAGCTTTCTTGGGTACAAGTCGGAGGGTCTATTCAATACAGAAAGGCAGACTAGG 6093
QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db 6094 CTCTCAGAAATTTTATATGATAGCACAGAGTCAGTTTACCTATGATGAACACAGCAGGA 6153
QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db 6154 GTCCTAAAGACAGATAACCTCCAGAGTGATGGTTTTATTGACCATTTAGATACAGSCAA 6213
QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db 6214 ATTGTCCTCTGATTGACAGGCAGATTTTCCGCTTTTAGTGAAGATGGGATGTAATGCA 6273
QY 321 ArgPheAspTyrAsnTyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsn 340
Db 6274 AGATTGACTATAGCTATGACAACACAGCTTTTCGAGTGACCAGCATGCAGGGTGTGATCAAT 6333
QY 341 GluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGluGln 360
Db 6334 GAAACGCCACTGCCCTATTGATCTGTATCAGTTTGATGATGACATTTCTGGCAAGTTGAGCAG 6393
QY 361 PheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMet 380
Db 6394 TTTGGAAAGTTTGGAGTTATATATATATGATATTAACAGATCATCTTCTACAGCTGTAATG 6453
QY 381 ThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePhe 400
Db 6454 ACCTATACGAAGCACCTTTGATGCTCATGGCCGCTATCAAGGAGATTCATATGAGATATTC 6513
QY 401 ArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLys 420

Db 6514 AGTGGCTCATGTACTGGATTACAATTCAATTCAGTATGATAAATGGTGGGTAACCAAGAGA 6573
QY 421 GluLeuLysValGlyProTyrAlaAsnThrArgTyrSerTyrGluTyrAspAlaAsp 440
Db 6574 GAGATTAAATAGGGCCCTTTGCCAACACACCACCAATATGCTTATGATATGATGTTGAT 6633
QY 441 GlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAspLeu 460
Db 6634 GGACAGCTCCAAACAGTTTACCTCAATGAAAGATAATGTGGCGGTACAACTACGATCTG 6693
QY 461 AsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyr 480
Db 6694 AATGGAACCTCCATTACTGAACCCCAAGTAACAGTGGCGTCTGACACCCCTTCGCTAT 6753
QY 481 AspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGly 500
Db 6754 GACCTGCGAGACAGAAATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGATGGT 6813
QY 501 PheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLys 520
Db 6814 TTCCTACGTCAAAGGGGCACGGAATCTTTGAATATAGTCCCAAGGGCTTCTAATCTGA 6873
QY 521 AlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgArg 540
Db 6874 GTTTACAGTAAAGGCGAGTGGCTGGACAGTGTATCTACCGTTATGACGGCCTGGGAAGCGT 6933
QY 541 Val 541
Db 6934 GTT 6936

RESULT 9
AAH14671
ID AAH14671 standard; cDNA; 3270 BP.
XX AC AAH14671;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:12354.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
XX
PS Claim 8; SEQ ID NO 12354; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 3270 BP; 867 A; 821 C; 852 G; 730 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.57e-209 Length: 3270
Score: 2159.00 Matches: 390
Percent Similarity: 89.53% Conservative: 72
Best Local Similarity: 75.58% Mismatches: 54
Query Match: 75.89% Indels: 0
DB: 4 Gaps: 0

US-10-029-020-14_COPY_1760_2300 (1-541) x AAH14671 (1-3270)
QY 26 HisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsnValThrLeuProIle 45
Db 3 CACGTTCTGGTGGCAGCGCTAATCCGACGGTGGCCAAAGAAACATGACTTGCCTGGC 62
QY 46 AspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGlnAlaArgGlyGlnVal 65
Db 63 GAGAACGGTCAAAACTTGGTGGATGGAGATTCGGAAGAGCAAGCCCAAGGAAAGTC 122
QY 66 ThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeuSerLeuAspPheAsp 85
Db 123 AATGTCTTTGGCCCAAGCTCAGGGTTAATGGCAGAAACCTCTTTCACTGACTTTGAT 182
QY 86 ArgValThrArgThrGluLysIleTyrAspAspHisArgLysPheThrLeuArgIleLeu 105
Db 183 CGAACACAAAGACAGAAAGATCTATGACGACCAACCGTAATTTCTACTGAGGATCGCC 242
QY 106 TyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArgLeuAsnGlyValAsn 125
Db 243 TAGACACGTCTGGGCACCCGACTCTCTGGTGGCCAGCAGCAAGCTGATGGCCGTCAT 302
QY 126 ValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGlyIleMetSerGluArg 145
Db 303 GTCACCTATTCCACAGGTCAAATTGCCAGCATCCAGCGAGGACCACTAGCGAGAAA 362
QY 146 MetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAlaAspGlyLysThrTrp 165
Db 363 GTAGATTATGACGACGAGGGGAGGATCGTGTCTCGGGTCTTTGCTGATGGTAAACATGG 422
QY 166 SerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSerGlnArgGlnTyrIle 185
Db 423 AGTTACACATATTTAGAAAGTCCATGGTTCTTCTGCTTATAGCCAGCGGAGTACATC 482
QY 186 PheGluPheAspLysAsnAspArgLeuSerSerValThrMetProAsnValAlaArgGln 205
Db 483 TTCGAATACGATATGTGGGACCGCCTGTCTGCCATCACCATGCCAGTGTGGCTCGCCAC 542
QY 206 ThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyrGlnProGluGly 225
Db 543 ACCATGCAGACCATCCGATCCATGGCTACTACCGCAACATATACAACCCCCCGGAAGC 602

QY 226 AsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeuHisThrPheTyrLeu 245
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 603 AACGCTCCATCATCACGGACTACAAACGAGGAGGCTGCTCTACAAACAGCTTTCCTTG 662

QY 246 GlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLysLeuAlaGluThrLeu 265
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 663 GGTACAAAGTCGGAGGCTCTTATTCAATAACAGAGGACAGACTAGGCTCTCAGAAATTTTA 722

QY 266 TyrAspThrLysValSerPheThrTyrAspGluThrAlaGlyMetLeuLysThrIle 285
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 723 TATGATAGCACAAAGAGTCAGTTTACCTATGATGATGAAACAGCAGGAGTCTCTAAAGACAGTA 782

QY 286 AsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGlnIleGlyProLeuIle 305
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 783 AACCTCCAGAGTGATGTTTATTTCACCATTAGATACAGGCAAAATGGTCCCTGATT 842

QY 306 AspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAlaArgPheAspTyrAsn 325
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 843 GACAGGCAGATTTCCGCTTTAGTGAAGATGGGATGTTAAATGCAAGATTTGACTATAGC 902

QY 326 TyrAspAsnSerPheArgValThrSerMetGlnAlaValIleAsnGluThrProLeuPro 345
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 903 TATGACAAACAGCTTTCGAGTGACCAAGCATGCGGTGTGATCAATGAAACGCCACTGCCT 962

QY 346 IleAspLeuTyrArgTyrAspValSerGlyLysThrGluGlnPheGlyLysPheGly 365
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 963 ATTGATCTGTATCAGTTTGATGACATTTCTGGCAAAAGTTGAGCAGTTTGGAAAGTTTGA 1022

QY 366 ValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaValMetThrHisThrLysHis 385
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1023 GTTATATATTATGATATTAAACAGATCATTTCTACAGCTGTAATGACCTATACGAAGCAC 1082

QY 386 PheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIlePheArgSerLeuMetTyr 405
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1083 TTGTGATGCTCATGGCCGATCAAGGAGATTCAATATGAGATATTTCAGGTCGCTCATGTAC 1142

QY 406 TrpMetThrValGlnTyrAspAsnMetGlyArgValValLysLysGluLeuLysValGly 425
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1143 TGGATTACAAATTCAGTATGATAACATGGTCCGGTAAACCAAGAGAGAGATTAAATAGGG 1202

QY 426 ProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAlaAspGlyGlnLeuGlnThr 445
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1203 CCCTTTGCCAACACCAACCAATATGCTTATGAATATGATGTGTGATGGACAGCTCCCAACA 1262

QY 446 ValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAspLeuAsnGlyAsnLeuHis 465
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1263 GTTTACCTCAATGAAAGAGATAATGTGGCGGTACAACTACGATCTGAATGGAAACCTCCAT 1322

QY 466 LeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArgTyrAspIleArgAspArg 485
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1323 TTACTGAACCAAGTAACAGTGGCGTCTGACACCCCTTCGCTATGACCTCGGAGACAGA 1382

QY 486 IleThrArgLeuGlyAspValGlnTyrLysMetAspGluAspGlyPheLeuArgGlnArg 505
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1383 ATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGATGGTTTCCTACGTCAAAGG 1442

QY 506 GlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIleLysAlaTyrAsnArgAla 525
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1443 GGCACGGAAATCTTTGAATATAGTCCCAAGGGCTTCTAACTCGAGTTTACAGTAAAGGC 1502

QY 526 GlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArgVal 541
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 1503 AGTGGCTGGACAGTATCTACCGTTATGACGGCCTGGGAAGGCGTGT 1550

RESULT 10
ACCT72051
ID ACC72051 standard; DNA; 9058 BP.
XX
AC ACC72051;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205A gene #SEQ ID 79.

XX Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.

XX Homo sapiens.
OS
XX WO2003029421-A2.
PN
XX 10-APR-2003.
PD
XX 02-OCT-2002; 2002WO-US031287.
PF
XX 03-OCT-2001; 2001US-0326526P.
PR
XX 14-MAY-2002; 2002US-00144194.
PR
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
PA
XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
PI
XX WPI; 2003-381623/36.
DR
XX P-PSDB; ABR58317.
DR
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.56e-206 Length: 9058
Score: 2136.50 Matches: 382
Percent Similarity: 87.06% Conservative: 89
Best Local Similarity: 70.61% Mismatches: 69
Query Match: 75.10% Indels: 1
DB: 7 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x ACC72051 (1-9058)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 4676 TACCAGCTCTGTAAATAATGGTACCTGAGGGTGATGTATGCTAATGGATGGGTATCAGC 4735

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 4736 TTCCACACAGCGAGCCCCCATGTCTCTAGCGGCGACCATCACCCCCACCATGGACGCTGCAAC 4795

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 4796 ATCTCCCTGGCTATGGAGATGGCTTAAACTCCATTGAGTGGCGCCCTAAGAAAGGACAG 4855

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 4856 ATTAAAGGCAAGTCACCATCTTTGGCAGGAAGCTCCGGTCCATGAGTAAAGAAATCTCTTG 4915

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspHisArgLysPhe 100
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
XX

Db 4916 TCCATTGACTATGATCGAAATATTTCGGACTGAAAGATCTATGATGACCACCGGAAGTTC 4975

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTyrSerProSerSerArg 120

Db 4976 ACCCTGAGGATCATTTATGACCAGGTGGGCGGCCCTTCCTCTGGCTGCCAGCAGCGG 5035

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140

Db 5036 CTGGCAGCTGTCAACGTGTCTACTTCTCAATGGGCGCCTGGCTGGCTTCAGCGTGG 5095

QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160

Db 5096 GCCATGAGCGAGAGGACAGACATCGACAAGCAAGCGCGCATCGTGTCCCGCATGTTTCGCT 5155

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuLeuHisSer 180

Db 5156 GACGGGAAAGTGTGGAGCTACTCCTACCTTGACAAGTCCATGGTTCCTCCTTCAGAGC 5215

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200

Db 5216 CAACGTCAGTATATATTTGAGTATGACTCCTCTGACCGCCTCCTTGCCGTCACCATGCC 5275

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220

Db 5276 AGCGTGGCCCCGGCACAGCATGTCCACACACACACCTCCATCGGTACATCCGTAATATTAC 5335

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Db 5336 AACC CGCCTGAAAGCAATGCTCGGTCTATCTTTGACTACAGTGATGACGGCCGCATCCTG 5395

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260

Db 5396 AAGACCTCCTTTTGGGCACCGGACCGCAGGTGTTCTACAAGTATGGAAACTTCCAAG 5455

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 5456 TTATCAGAGATTGTCTACGACAGTACCGCGCTCACCTTCGGGTATGACGAGACCACTGGT 5515

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db 5516 GTCTTGAAGATGGTCAACCTCCAAAGTGGGGGCTTCTCCTGCACCATCAGGTACCGAAG 5575

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320

Db 5576 ATTGCCCCCTGGTGGACAAGCAGATCTACAGTTCTCCGAGGAAGGCATGGTCAATGCC 5635

QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339

Db 5636 AGGTTTGACTACACCTATCATGACAACAGCTTCGCGCATCGCAAGCATCAAGCCCGTCATA 5695

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu 359

Db 5696 AGTGAGACTCCCTCCCGTTTGACCTTACCGCTATGATGAGATTCTGGCAAGGTGGAA 5755

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379

Db 5756 CACTTTGGTAAGTTGGAGTCATCTATTATGACATCAACAGATCATCACCACCTGCCGTG 5815

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399

Db 5816 ATGACCCCTCAGCAAAACACTTCGACACCCATGGGCGGATCAAGGAGGTCCAGTATGAGATG 5875

QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValLys 419

Db 5876 TTCCGTCCTCATGTACTGTGATGACGGTGCAATATGACAGCATGGCAGGTGATCAAG 5935

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439

Db 5936 AGGGAGCTAAAACTGGGCCCCCTATGCCAATACCACGAAGTACACCTATGACTACGATGG 5995

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459

Db 5996 GACGGGCAGCTCCAGAGCGGTGGCGTCAATGACCGCCGACCTGGCGCTACAGCTATGAC 6055

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479

Db 6056 CTTAATGGGAATCTCCACTTACTGACCCAGGCAACAGTGTGCGCCTCATGCCCTTGGC 6115

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499

Db 6116 TATGACCTCCGGGATCGGATAACCACTCGGGGATGTGCAGTACAAAATTGACGACGAT 6175

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519

Db 6176 GGCTATCTGTGCCAGAGAGGTCTGACATCTTCGAATACAATTCCAAGGGCCTCCTAACA 6235

QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539

Db 6236 AGAGCCTACAACAAGGCCAGCGGTGGAGTGTCCAGTACCGCTATGATGGCGTAGGACGG 6295

QY 540 Arg 540

Db 6296 CGG 6298

RESULT 11

ACC72052

ID ACC72052 standard; DNA; 9695 BP.

XX

AC ACC72052;

XX

DT 08-JUL-2003 (first entry)

XX

DE BCU0205B gene #SEQ ID 81.

XX

KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated; drug discovery; clinical medicine; forensic medicine; gene; chromosome 5q33.3; ds.

KW

XX Homo sapiens.

OS

XX WO2003029421-A2.

PN

XX PD 10-APR-2003.

XX

PF 02-OCT-2002; 2002WO-US031287.

XX

PR 03-OCT-2001; 2001US-0326526P.

PR 14-MAY-2002; 2002US-00144194.

XX

PA (ORIG-) ORIGENE TECHNOLOGIES INC.

XX

PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;

XX

DR WPI; 2003-381623/36.

DR P-PSDB; ABR58318.

XX

PT New isolated human differentially-regulated breast cancer polynucleotide and polypeptide, useful for diagnosing, staging, prognosticating, preventing and/or treating diseases and conditions relating to breast cancer.

PT

XX

PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.

XX

CC The invention relates to isolated polynucleotides which are differentially-regulated in breast cancer. The methods and compositions of the present invention are useful for detecting, diagnosing, staging, monitoring, prognosticating, preventing and/or treating diseases and conditions relating to breast cancer, and may be used in gene therapy or antisense therapy. They can also be used in research, drug discovery, clinical medicine and forensic medicine. Sequences given in records ACC72012-ACC72074 represent polynucleotides of the invention that are differentially-regulated in breast cancer. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
Db 6441 CACTTGGTAAGTTTGGAGTCATCTATTATGACATCAACCCAGATCATCACCACGTGCCGTG 6500
QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db 6501 ATGACCCTCAGCAAAACACTTCGACACCCATGGCGGATCAAGAGGTCCAGTATGAGATG 6560
QY 400 PheArgSerLeuMetTyrTyrMetThrValGlnTyrAspAsnMetGlyArgValVallys 419
Db 6561 TTCCGGTCCCTCATGTACTGTGATGACGGTGCAATATGACAGCATGGCAGGTGATCAAG 6620
QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db 6621 AGGGAGCTAAACTGGGGCCCTATGCCAATACACAGAGTACACCTATGACTACGATGGG 6680
QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp 459
Db 6681 GACGGGCAGCTCCAGAGCGTGGCCGTCAATGACCGCCCGACCTGGCGCTACAGCTATGAC 6740
QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db 6741 CTTAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGGCCTCATGGCCTTGCGC 6800
QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db 6801 TATGACCTCCGGATCGGATAACACAGACTCGGGGATGTGCAGTACAAAATTGACGACGAT 6860
QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
Db 6861 GGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATACAAATTCGAAGGGCCCTCTAACA 6920
QY 520 LysAlaTyrAsnArgAlaGlySerTyrSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 6921 AGAGCCTACAACAAGGCCAGCGGGTGGAGTGTCCAGTACCGCTATGATGGCGTAGGACGG 6980
QY 540 Arg 540
Db 6981 CGG 6983
RESULT 13
ADB32028
ID ADB32028 standard; cDNA; 9729 BP.
XX
AC ADB32028;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human FCTR3f cDNA.
XX
KW Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX
OS Homo sapiens.
XX
PN US2003087816-A1.
XX
PD 08-MAY-2003.
XX
PF 05-MAR-2001; 2001US-00800198.
XX

03-MAR-2000; 2000US-0186592P.
PR
XX
PA (VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RAST/) RASTELLI L.
XX
PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2003-625633/59.
XX
DR P-PSDB; ADB32029.
PT
PT New FCTRX polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTRX-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
XX
PS Claim 9; Page 34-37; 155pp; English.
XX
CC The invention relates to FCTRX polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTRX polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTRX polypeptide of the invention.
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.92e-206 Length: 9729
Score: 2136.50 Matches: 382
Percent Similarity: 87.06% Conservative: 89
Best Local Similarity: 70.61% Mismatches: 69
Query Match: 75.10% Indels: 1
DB: Gaps: 1
US-10-029-020-14_COPY_1760_2300 (1-541) x ADB32028 (1-9729)
QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5361 TACCAGCTCTGTAATAATGGTACCCCTGAGGGTGATGTATGCTAATGGGATGGGTATCAGC 5420
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5421 TTCCACAGCGAGCCCCATGTCTCAGCGGGCACCATCACCCCCACCATGGACGCTGCAAC 5480
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTyrArgGlnArgLysGluGln 60
Db 5481 ATCTCCCTGCCTATGGAGAATGGCTTAACCTCCATTGAGTGGCGCTTAAGAAAGGAACAG 5540
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db 5541 ATTAAGGCAAGTCACCATCTTTGGCAGGAAGCTCCGGGTCCATGGGAAGAAATCTCTTG 5600
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspHisArgLysPhe 100

Db	5601	ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerArg	120
Db	5661	ACCCTGAGGATCATTTATGACCAGGTGGGCGCCCTTCTCTGGCTGCCAGACGCGG	5720
QY	121	LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly	140
Db	5721	CTGGCAGCTGTCAACGTGTCTATCTTCAATGGGCGCCTGGCTGGCTTCAGCGTGGG	5780
QY	141	IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla	160
Db	5781	GCCATGACGAGAGGACAGACATCGACAAGCAAGCGCGCATCGTGTCCCAGCATGTCGCT	5840
QY	161	AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer	180
Db	5841	GACGGAAAGTGTGGAGCTACTCTACCTTGACAAGTCCATGGTCTCTCTGCTTCAGAGC	5900
QY	181	GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro	200
Db	5901	CAACGTGAGTATATTGAGTATGACTCTCTGACCGCTCTTGGCGTCCACCATGCC	5960
QY	201	AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr	220
Db	5961	AGCGTGGCCCGGCACAGCATGTCCACACACACCTCCATCGCTACATCCCGTAATTTAC	6020
QY	221	GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu	240
Db	6021	AACCCGCTGAAAGCATGCTTCGGTCACTTTTGACTACAGTATGACGCGCGCATCCTG	6080
QY	241	HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys	260
Db	6081	AAGACCTCCTTTTGGGACCGGACGCCAGGTGTCTACAAGTATGGGAACTCTCCAAG	6140
QY	261	LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly	280
Db	6141	TTATCAGAGATTGTCTACGACAGTACCGCCGTCACTTCGGGTATGACGAGACCACTGGT	6200
QY	281	MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln	300
Db	6201	GTCTTGAAGATGGTCAACCTCCAAAGTGGGGGCTTCTCCTGCACCATCAGTACCGAAG	6260
QY	301	IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla	320
Db	6261	ATTGGCCCCCTGGTGGACAAGCAGATCTACAGGTCTCCGAGGAAGCATGGTCAATGCC	6320
QY	321	ArgPheAspTyrAsnTyr--AspAsnSerPheArgValThrSerMetGlnAlaValIle	339
Db	6321	AGGTTTGACTACACCTATCATGACAACAGCTTCGCGATCGCAAGCATCAAGCCCCGTATA	6380
QY	340	AsnGluThrProLeuProIleAspLeuTyrArgTyrAspAspValSerGlyLysThrGlu	359
Db	6381	AGTGAGACTCCCTCCCGTTGACCTCTACCGCTATGATGAGATTCTTGCAAGGTGGAA	6440
QY	360	GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal	379
Db	6441	CACTTTGGTAAGTTTGGAGTCATCTATTATGACATCAACCAGATCATCACCCTGCCGTG	6500
QY	380	MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle	399
Db	6501	ATGACCCCTCAGCAAAACACTTCGACACCCATGGCGGATCAAGGAGGTCCAGTATGAGATG	6560
QY	400	PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLys	419
Db	6561	TTCCGGTCCCTCATGTACTGGATGACGGTGCATATGACAGCATGGCGAGGTGATCAAG	6620
QY	420	LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla	439
Db	6621	AGGGAGCTAAAACTGGGGCCCTATGCCAATACCACGAAGTACACCTATGACTACGATGGG	6680
QY	440	AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTyrArgTyrSerTyrAsp	459

Db	6681	GACGGCAGCTCCAGAGCGGTGGCCGTCAATGACCCGCCGACCTGGCGCTACAGCTATGAC	674
QY	460	LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg	479
Db	6741	CTTAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTCGCCTCATGCCCCTTGGCG	6800
QY	480	TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp	499
Db	6801	TATGACCTCCGGATCGGATAACCAGACTCGGGGATGTGCAGTACAAAATTGACGACGAT	6860
QY	500	GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle	519
Db	6861	GGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATACAATTCCAAGGGCCTCCTAACA	6920
QY	520	LysAlaTyrAsnArgAlaGlySerTipSerValArgTyrArgTyrAspGlyLeuGlyArg	539
Db	6921	AGAGCCTACAACAAGGCCAGCGGGTGGAGTGTCCAGTACCGCTATGATGGCGTAGGACGG	6980
QY	540	Arg 540	
Db	6981	CGG 6983	
RESULT 14			
AAS14085			
ID	AAS14085 standard; DNA; 9826 BP.		
XX			
AC	AAS14085;		
XX			
DT	18-DEC-2001 (first entry)		
XX			
DE	Human FCTR3b DNA sequence.		
XX			
KW	Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;		
KW	astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;		
KW	neurological disorder; neurodegenerative disorders; nerve trauma;		
KW	familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;		
KW	demyelinating Gardner syndrome; familial myelodysplastic syndrome;		
KW	mental health condition; immunological disorder; allergy; infertility;		
KW	bronchial asthma; Avellino type eosinophilia; lung disease; deafness;		
KW	reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;		
KW	desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;		
KW	gastric disorders; pancreatic disease; Schistosoma mansoni infection;		
KW	Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;		
KW	Corneal dystrophy-Greonouw type I; Corneal dystrophy-lattice type I;		
KW	Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;		
KW	antiallergic; antiasthmatic; antiinfertility; antiinflammatory;		
KW	antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;		
KW	gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;		
KW	gene therapy; FCTR3b; neurestin-like protein.		
XX			
OS	Homo sapiens.		
XX			
FH	Key		
FT	5'UTR	Location/Qualifiers	
FT		1..279	
FT		/*tag= b	
FT	CDS	280..8481	
FT		/*tag= a	
FT		/product= "Human FCTR3b"	
FT	3'UTR	8482..9826	
FT		/*tag= c	
XX			
PN	WO200166747-A2.		
XX			
PN	13-SEP-2001.		
PD			
XX			
PF	05-MAR-2001; 2001WO-US007160.		
XX			
PR	03-MAR-2000; 2000US-0186592P.		
PR	03-MAR-2000; 2000US-0186718P.		
PR	06-MAR-2000; 2000US-0187293P.		
PR	06-MAR-2000; 2000US-0187294P.		
PR	17-MAR-2000; 2000US-0190400P.		
PR	07-APR-2000; 2000US-0196018P.		


```
PR 03-JAN-2001; 2001US-0259548P.
XX
PA (CURA-) CURAGEN CORP.
XX Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
PI
XX
DR WPI; 2001-596837/67.
DR P-PSDB; AAU08680.
XX
PT Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
PS Claim 9; Page 33-35; 215pp; English.
XX
CC The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
CC nucleic acids encoding them. These sequences are useful for the treatment
CC or prevention of numerous disorders including myelogenous leukaemia,
CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
CC disorders, thrombocytopaenia, neurological disorders, neurodegenerative
CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avelilno type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greonow type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC FCTR3b, a neurestin-like protein
XX
SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.97e-206 Length: 9826
Score: 2136.50 Matches: 382
Percent Similarity: 87.06% Conservative: 89
Best Local Similarity: 70.61% Mismatches: 69
Query Match: 75.10% Indels: 1
DB: 5 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x AAS14085 (1-9826)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuLeuAlaAsnGlyMetGluValAla 20
Db TACCAGCTCTGTAATAATGGTACCTGAGGTGATGTATGCTAATGGGATGGGTATCAGC 5517

QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db TTCACAGCGAGCCCCATGTCTTAGCGGCACCATCACCCCCACCATTTGGACCGTGCAC 5577

QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db ATCTCCCTGCCCTATGGAGAAATGGCTTAAACTCCATTGAGTGGCGCCTAAGAAAGAACAG 5637

QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeuLeu 80
Db ATTAAGGCAAAAGTCACCATCTTTGGCAGGAAGCTCCGGGTCCATGGAAGAAATCTCTTG 5697

QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db TCCATTGACTATGATCGAAATATTCCGACTGAAAGATCTATGATGACCAACCGGAAGTTC 5757

QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db ACCCTGAGGATCATTTATGACCAGGTGGCGGCCCTTCTCTGGCTGCCACGACGCGG 5817

QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db CTGGCAGCTGTCAACGTGTACATACTTCTTCAATGGGCGCCTGGCTGGGCTTCAGCGTGGG 5877
```

```
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db GCCATGAGCGAGAGGACAGACATCGACAAGCAAGCGCGCATCGTGTCCCGCATGTCGCT 5937

QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db GACGGGAAAGTGTGGAGCTACTCTACCTTGACAAGTCCATGGTCTCTCTGCTTCAGAGC 5997

QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db CAACGTCACTATATATTGAGTATGACTCTCTGACCGCTCTCTGCGCTACCATGCCC 6057

QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db AGCGTGGCCCGGCACAGCATGTCCACACACACCTCCATCGGCTACATCCGTAATATTAC 6117

QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240
Db AACCCGCTGAAAGCAATGCTTCGGTCTCTTTGACTACAGTATGACGCGCGCATCTCTG 6177

QY 241 HisThrPheTyrLeuGlyThrGlyArgArgValIleTyrLysTyrGlyLysLeuSerLys 260
Db AAGACCTCTCTTTTGGGCACCGGACCGCGAGGTGTCTACAAGTATGGGAAACTCTCCAAG 6237

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280
Db TTATCAGAGATTGTCTACGACAGTACCGCCGTCACCTTCGGGTATGACGAGACCACTGGT 6297

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300
Db GTCTTGAAGATGTCAACCTCCAAAGTGGGGCTTCTCTGCACCATCAGGTACCGGAAG 6357

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320
Db ATTGGCCCCCTGGTGGACAAGCAGATCTACAGGTCTCTCGAGGAAGGATGGTCAATGCC 6417

QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339
Db AGGTTTGACTACACCTATCATGACAACACAGCTTCCGCGCATCGCAAGCATCAAGCCCGTCATA 6477

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359
Db AGTGAGACTCCCTCCCGTTGACCTCTACCGCTATGATGAGATTTCTGGCAAGGTGGAA 6537

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379
Db CACTTTGGTAAGTTTGGAGTCATCTATTATGACATCAACAGATCATCACCACCTGCCGTG 6597

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399
Db ATGACCTCAGCAAAACACTTCGACACCCATGGCGGATCAAGGAGGTCCAGTATGAGATG 6657

QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValValLys 419
Db TTCCGCTCCCTCATGTACTGGATGACGGTGCAATATGACAGCATGGGCGGATGATCAAG 6717

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439
Db AGGAGCTAAACTGGGGCCCTATGCCAATACCACGAAGTACACCTATGACTACGATGGG 6777

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459
Db GACGGGAGCTCCAGAGCGGTGGCGCTCAATGACCGCCCGACCTGGCGCTACAGCTATGAC 6837

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479
Db CTTAATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGCGCTCATGCCCTTGGCG 6897

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499
Db TATGACCTCCGGGATCGGATAACCAAGACTCGGGGATGTGCAGTACAAATGACGACGAT 6957

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519
```

Db 6958 GGCTATCTGTGCCAGAGAGGGTCTGACATCTTCGAATACAATTCCAAGGGCCTCTCTAACA 7017
QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539
Db 7018 AGAGCCTACACACAGGCCAGCGGGTGGAGTGTCAGTACCGCTATGATGGCGTAGGACGG 7077
QY 540 Arg 540
Db 7078 CGG 7080
RESULT 15
ADB32023
ID ADB32023 standard; cDNA; 9826 BP.
XX AC ADB32023;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human FCTR3b cDNA.
XX
KW Human; gene; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX
OS Homo sapiens.
XX
PN US2003087816-A1.
XX
PD 08-MAY-2003.
XX
PF 05-MAR-2001; 2001US-00800198.
XX
PR 03-MAR-2000; 2000US-0186592P.
XX
PA (VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RAST/) RASTELLI L.
PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K,
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2003-625633/59.
XX P-PSDB; ADB32024.
XX
PT New FCTRX polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTRX-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
XX
PS Claim 9; Page 29-32; 155pp; English.
XX
CC The invention relates to FCTRX polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTRX polypeptide, such as colorectal cancer,

CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanoma, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTRX polypeptide of the invention.
XX
SQ Sequence 9826 BP; 2567 A; 2645 C; 2476 G; 2138 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.97e-206 Length: 9826
Score: 2136.50 Matches: 382
Percent Similarity: 87.06% Conservative: 89
Best Local Similarity: 70.61% Mismatches: 69
Query Match: 75.10% Indels: 1
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1760_2300 (1-541) x ADB32023 (1-9826)

QY 1 TyrTyrIleGlyAlaAspGlySerLeuArgLeuLeuAlaAsnGlyMetGluValAla 20
Db 5458 TACCAGCTCTGTAATAATGGTACCTGAGGTGATGTATGCTAATGGGATGGGTATCAGC 5517
QY 21 LeuGlnThrGluProHisLeuLeuAlaGlyThrValAsnProThrValGlyLysArgAsn 40
Db 5518 TTCACAGCGAGCCCCCATGTCTCTAGCGGCGCACCATCACCCCCCATTTGGACGCTGCAAC 5577
QY 41 ValThrLeuProIleAspAsnGlyLeuAsnLeuValGluTrpArgGlnArgLysGluGln 60
Db 5578 ATCTCCCTGCCTATGGAGATGGCTTAACTCCATGAGTGGCGCTTAAGAAAGGAACAG 5637
QY 61 AlaArgGlyGlnValThrValPheGlyArgArgLeuArgValHisAsnArgAsnLeu 80
Db 5638 ATTAAAGGCAAGTCACCATCTTTGGCAGGAAGTCCGGGTCCATGGAGAAATCTCTTG 5697
QY 81 SerLeuAspPheAspArgValThrArgThrGluLysIleTyrAspAspHisArgLysPhe 100
Db 5698 TCCATTGACTATGATCGAAATATTCCGAGCTGAAAGATCTATGATGACCAACCGAAGTTC 5757
QY 101 ThrLeuArgIleLeuTyrAspGlnAlaGlyArgProSerLeuTrpSerProSerSerArg 120
Db 5758 ACCCTGAGGATCATTTATGACCAGGTGGCGGCCCTTCTCTGGTGGCCAGCAGCGGG 5817
QY 121 LeuAsnGlyValAsnValThrTyrSerProGlyGlyTyrIleAlaGlyIleGlnArgGly 140
Db 5818 CTGGCAGCTGTCAACGTGTACATCTTCAATGGCGCCTGGCTGGCTTCAGCGTGGG 5877
QY 141 IleMetSerGluArgMetGluTyrAspGlnAlaGlyArgIleThrSerArgIlePheAla 160
Db 5878 GCCATGACGAGAGGACAGACATCGACACCAAGCCCGCATCGTCTCCCGCATGTTTCGCT 5937
QY 161 AspGlyLysThrTrpSerTyrThrTyrLeuGluLysSerMetValLeuLeuHisSer 180
Db 5938 GACGGGAAAGTGTGGAGCTACTCTACCTTGACAAGTCCATGGTCTCTCTGCTTCAGAGC 5997
QY 181 GlnArgGlnTyrIlePheGluPheAspLysAsnAspArgLeuSerSerValThrMetPro 200
Db 5998 CAACGTCAGTATATTTGAGTATGACTCTCTGACCGCCTCTTGGCTGCCGTCACCATGCC 6057
QY 201 AsnValAlaArgGlnThrLeuGluThrIleArgSerValGlyTyrTyrArgAsnIleTyr 220
Db 6058 AGCGTGGCCCGCACAGCATGTCCACACACACCTCCATCGGCTACATCCGTAATATTAC 6117
QY 221 GlnProProGluGlyAsnAlaSerValIleGlnAspPheThrGluAspGlyHisLeuLeu 240

Mon Aug 16 09:01:10 2004

Db 6118 AACCCGCTGAAAGCAATGCTTCGGTCATCTTTGACTACAGTGATGACGGCGGCATCCTG 6177

QY 241 HisThrPheTyrLeuGlyThrGlyArgValIleTyrLysTyrGlyLysLeuSerLys 260

Db 6178 AAGACCTCTTTTGGGCACCGGAGCCAGGTGTTCTACAGTATGGGAAACTCTCCAAG 6237

QY 261 LeuAlaGluThrLeuTyrAspThrThrLysValSerPheThrTyrAspGluThrAlaGly 280

Db 6238 TTATCAGAGATTGCTACGACAGTACCGCCGTCACTTCGGGTATGACGAGACCACTGGT 6297

QY 281 MetLeuLysThrIleAsnLeuGlnAsnGluGlyPheThrCysThrIleArgTyrArgGln 300

Db 6298 GTCTTGAGATGGTCAACCTCCAAGTGGGGCTTCTCTGCACCATCAGGTACCCGGAAG 6357

QY 301 IleGlyProLeuIleAspArgGlnIlePheArgPheThrGluGluGlyMetValAsnAla 320

Db 6358 ATTGCCCCCTGGTGGACAAGCAGATCTACAGGTTCTCCGAGGAAGGCATGGTCAATGCC 6417

QY 321 ArgPheAspTyrAsnTyr---AspAsnSerPheArgValThrSerMetGlnAlaValIle 339

Db 6418 AGGTTTGACTACACCTATCATGACAACAGCTTCGGCATCGCAAGCATCAAGCCCGTCATA 6477

QY 340 AsnGluThrProLeuProIleAspLeuTyrArgTyrAspValSerGlyLysThrGlu 359

Db 6478 AGTGAGACTCCCTCCCGTTGACCTCTACCGCTATGATGAGATTTCTGGCAAGGTGGAA 6537

QY 360 GlnPheGlyLysPheGlyValIleTyrTyrAspIleAsnGlnIleIleThrThrAlaVal 379

Db 6538 CACTTTGGTAAGTTGGAGTCATCTATTATGACATCAACCCAGATCATCACCTGCCGTG 6597

QY 380 MetThrHisThrLysHisPheAspAlaTyrGlyArgMetLysGluValGlnTyrGluIle 399

Db 6598 ATGACCTCAGCAAAACACTTCGACACCCATGGGCGGATCAAGGAGGTCCAGTATGAGATG 6657

QY 400 PheArgSerLeuMetTyrTrpMetThrValGlnTyrAspAsnMetGlyArgValVallLys 419

Db 6658 TTCCGGTCCCTCATGTACTGGATGACGGTGCAATATGACAGCATGGGCGGTGATCAAG 6717

QY 420 LysGluLeuLysValGlyProTyrAlaAsnThrThrArgTyrSerTyrGluTyrAspAla 439

Db 6718 AGGGAGCTAAAACTGGGGCCCTATGCCAATACCAAGAGTACACCTATGACTACGATGGG 6777

QY 440 AspGlyGlnLeuGlnThrValSerIleAsnAspLysProLeuTrpArgTyrSerTyrAsp 459

Db 6778 GACGGCAGCTCCAGAGCGTGGCGTCAATGACCCCGACCTGGCGCTACAGCTATGAC 6837

QY 460 LeuAsnGlyAsnLeuHisLeuLeuSerProGlyAsnSerAlaArgLeuThrProLeuArg 479

Db 6838 CTTATGGGAATCTCCACTTACTGAACCCAGGCAACAGTGTGGCCTCATGCCCTTGCGC 6897

QY 480 TyrAspIleArgAspArgIleThrArgLeuGlyAspValGlnTyrLysMetAspGluAsp 499

Db 6898 TATGACCTCCGGATCGGATAACCCAGACTCGGGGATGTGCAGTACAAAATTGACGCGAT 6957

QY 500 GlyPheLeuArgGlnArgGlyGlyAspIlePheGluTyrAsnSerAlaGlyLeuLeuIle 519

Db 6958 GGCTATCTGTCCAGAGAGGGTCTGACATCTTCGAATACAAATTCGAAGGGCCTCCTAACA 7017

QY 520 LysAlaTyrAsnArgAlaGlySerTrpSerValArgTyrArgTyrAspGlyLeuGlyArg 539

Db 7018 AGAGCCTACAAAGGCCAGCGGGTGGAGTGTCCAGTACCGCTATGATGGCGTAGGACGG 7077

QY 540 Arg 540

Db 7078 CGG 7080

Search completed: August 14, 2004, 02:33:50
Job time : 1049.68 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 523.097 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750
Perfect score: 1384
Sequence: 1 KINRIRQVTTSGEISLVAGA.....SSKDDVTITNLSASGAFYT 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn_1_1_2156@runat_06082004_112218_29331
-NCFU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq: *
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq: *
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq: *
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq: *
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq: *
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq: *
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq: *
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq: *
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq: *
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq: *
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq: *
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq: *
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2: *
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq: *
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq: *
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq: *
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq: *
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq: *
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
§				

1	1384	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
2	1384	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
3	1384	100.0	8438	13	US-10-042-865-1	Sequence 1, Appli
4	1379	99.6	8355	13	US-10-383-201-55	Sequence 55, Appl
5	935.5	67.6	8473	17	US-10-038-854-39	Sequence 39, Appl
6	935.5	67.6	8487	17	US-10-038-854-41	Sequence 41, Appl
7	935.5	67.6	8645	17	US-10-038-854-37	Sequence 37, Appl
8	935.5	67.6	8675	17	US-10-038-854-35	Sequence 35, Appl
9	898	64.9	12880	16	US-10-295-027-927	Sequence 927, App
10	865.5	62.5	8689	9	US-09-808-602-78	Sequence 78, Appl
11	865.5	62.5	8689	10	US-09-800-198-66	Sequence 66, Appl
12	864.5	62.5	8797	9	US-09-808-602-74	Sequence 74, Appl
13	864.5	62.5	8797	9	US-09-808-602-77	Sequence 77, Appl
14	864.5	62.5	8797	10	US-09-800-198-62	Sequence 62, Appl
15	864.5	62.5	8797	10	US-09-800-198-65	Sequence 65, Appl
16	863.5	62.4	6560	9	US-09-808-602-76	Sequence 76, Appl
17	863.5	62.4	6560	10	US-09-800-198-64	Sequence 64, Appl
18	863.5	62.4	9058	16	US-10-144-194A-79	Sequence 79, Appl
19	863.5	62.4	9695	16	US-10-144-194A-81	Sequence 81, Appl
20	863.5	62.4	9729	9	US-09-808-602-12	Sequence 12, Appl
21	863.5	62.4	9729	10	US-09-800-198-12	Sequence 12, Appl
22	863.5	62.4	9826	9	US-09-808-602-7	Sequence 7, Appli
23	863.5	62.4	9826	10	US-09-800-198-7	Sequence 7, Appli
24	862.5	62.3	8409	9	US-09-808-602-79	Sequence 79, Appl
25	862.5	62.3	8409	10	US-09-800-198-67	Sequence 67, Appl
26	861.5	62.2	8575	13	US-10-072-012-143	Sequence 143, App
c 27	389	28.1	3310	13	US-09-823-245A-482	Sequence 482, App
28	270.5	19.5	3614	13	US-10-342-887-1743	Sequence 1743, Ap
29	270.5	19.5	3614	13	US-10-172-118-1743	Sequence 1743, Ap
30	160	11.6	450	15	US-10-029-386-16526	Sequence 16526, A
31	160	11.6	555	15	US-10-029-386-2826	Sequence 2826, Ap
c 32	159	11.5	457	16	US-10-027-632-183606	Sequence 183606,
c 33	159	11.5	457	16	US-10-027-632-183606	Sequence 183606,
34	133.5	9.6	182	15	US-10-029-386-16081	Sequence 16081, A
35	133.5	9.6	519	15	US-10-029-386-2381	Sequence 2381, Ap
36	131.5	9.5	3750	15	US-10-156-761-22	Sequence 22, Appl
c 37	131.5	9.5	9025608	15	US-10-156-761-1	Sequence 1, Appli
38	119.5	8.6	4716	13	US-10-282-122A-33231	Sequence 33231, A
39	117.5	8.5	3957	16	US-10-369-493-34930	Sequence 34930, A
40	115	8.3	31960	15	US-10-114-170-11	Sequence 11, Appl
c 41	111.5	8.1	1235	16	US-10-398-221-1343	Sequence 1343, Ap
42	111.5	8.1	4533	16	US-10-418-861B-54	Sequence 54, Appl
c 43	110.5	8.0	5394	16	US-10-398-221-3837	Sequence 3837, Ap
44	109.5	7.9	6392	16	US-10-398-221-689	Sequence 689, App
45	109.5	7.9	6504	16	US-10-398-221-2219	Sequence 2219, Ap

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

Db	5162	ACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGCATGTCCAGGTAGAG	5221
Qy	241	ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr	260
Db	5222	ACCTCCAGCAAGGATGATGTCAACATAAACCAACCACTGCTGCCTCAGGCGCCTTCTAC	5281
Qy	261	Thr 261	
Db	5282	ACA 5284	
RESULT 2			
US-10-029-020-13			
; Sequence 13, Application US/10029020			
; Publication No. US20040033971A1			
; GENERAL INFORMATION:			
; APPLICANT: Gangolli et al.			
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same			
; FILE REFERENCE: 21402-225			
; CURRENT APPLICATION NUMBER: US/10/029,020			
; PRIOR FILING DATE: 2001-12-19			
; PRIOR APPLICATION NUMBER: 60/256,704			
; PRIOR FILING DATE: 2000-12-19			
; PRIOR APPLICATION NUMBER: 60/311,590			
; PRIOR FILING DATE: 2001-08-10			
; PRIOR APPLICATION NUMBER: 60/257,314			
; PRIOR FILING DATE: 2000-12-20			
; PRIOR APPLICATION NUMBER: 60/311,613			
; PRIOR FILING DATE: 2001-08-10			
; PRIOR APPLICATION NUMBER: 60/315,617			
; PRIOR FILING DATE: 2001-08-29			
; PRIOR APPLICATION NUMBER: 60/307,506			
; PRIOR FILING DATE: 2001-07-24			
; PRIOR APPLICATION NUMBER: 60/322,358			
; PRIOR FILING DATE: 2001-09-14			
; PRIOR APPLICATION NUMBER: 60/294,075			
; PRIOR FILING DATE: 2001-05-29			
; PRIOR APPLICATION NUMBER: 60/288,153			
; PRIOR FILING DATE: 2001-05-02			
; NUMBER OF SEQ ID NOS: 190			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 13			
; LENGTH: 8354			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-029-020-13			
Alignment Scores:			
Pred. No.:	1.99e-170	Length:	8354
Score:	1384.00	Matches:	261
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	13	Gaps:	0
US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-029-020-13 (1-8354)			
Qy	1	LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla	20
Db	4502	AAGATCAACCGCATCAGGCAGGTCAACCACTAGTGAGAGATCTCACTCGTTGCTGGGGCC	4561
Qy	21	ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly	40
Db	4562	CCCAGTGGCTGTGACTGTAAATAATGATGCCAATCTTCTTCTGGAGACGATGGT	4621
Qy	41	TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu	60
Db	4622	TATGCCAAGGATGCAAAGTTAAATACCCCATCTTCCCTGGTGTGTGTGCTGATGGGGAG	4681
Qy	61	LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe	80
Db	4682	CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGAAGAACAAAGCCTTTC	4741
Qy	81	LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe	100
Db	4742	CTCAACACCCAGAACATGTATGAGCTGTCTTACCAATTGACAGGAGCTCTATCTGTTT	4801
Qy	101	AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn	120
Db	4802	GATACCAACCGGCAAGCACCTGTACACCAAGCCCTGCCCCACAGGAGACTACCTGTACAAC	4861
Qy	121	PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal	140
Db	4862	TTCACCTACACTGGGGACGGCGACATCACACTCATCACAGACAAATGGCAACATGGTA	4921
Qy	141	AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal	160
Db	4922	AATGTCCGCGAGACTCTACTGGGATGCCCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGT	4981
Qy	161	TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu	180
Db	4982	TACTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGACCAACAAAGCAATGAAAC	5041
Qy	181	LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn	200
Db	5042	TTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAACCAAGCAATGAAAC	5101
Qy	201	GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro	220
Db	5102	GGATGGACAACATTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCCT	5161
Qy	221	ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu	240

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProilleAspGlnGluLeuTyrLeuPhe 100
Db 4742 CTC AACACCCAGAACATGTATGAGCTGTCTTACCAATTGACGAGGCTCTATCTGTTT 4801

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4802 GATACCAACCGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 4861

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4862 TTCACCTACACTGGGACGGCGACATCACACTCATCACAGACAACAATGGCAACATGGTA 4921

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4922 AATGTCGCGGAGACTCTACTGGGATGCCCTCTGGCTGGTCCAGATGGCCAGGTG 4981

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4982 TACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACCACACAAGGACACGAG 5041

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5042 TTGGCCATGATGACATACCATGGCAATTCGGCCTTCTGGCAACCAAAAGCAATGAAAC 5101

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5102 GGATGGACAACATTTTATGAGTAGACAGACTTTGGCGCCTGACAAATGTGACCTTCCCT 5161

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5162 ACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGCATGTCCAGGTAGAG 5221

QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5222 ACCTCCAGCAAGGATGATGTACCATAACCAACCACTGTCTGCCTCAGGCGCCTTCTAC 5281

QY 261 Thr 261
Db 5282 ACA 5284

RESULT 3

US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchexnev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik

Alignment Scores:
Pred. No.: 2.02e-170 Length: 8438
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-042-865-1 (1-8438)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4508 AAGATCAACCGCATCAGGCAGGTCAACCACTAGTGAGAGATCTCACTCGTTGCTGGGCC 4567

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4568 CCCAGTGGCTGTGACTGTAAATAATGATGCCAATGTGATGTTGTTTCTGGAGACGATGGT 4627

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4628 TATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGAG 4687

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db 4688 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAAACAAGCCTTTC 4747

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProilleAspGlnGluLeuTyrLeuPhe 100
Db 4748 CTC AACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACGAGGACTCTATCTGTTT 4807

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4808 GATACCAACCGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 4867

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4868 TTCACCTACACTGGGACGGCGACATCACACTCATCACAGACAACAATGGCAACATGGTA 4927

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4928 AATGTCGCGGAGACTCTACTGGGATGCCCTCTGGCTGGTGTGCCAGATGGCCAGGTG 4987

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4988 TACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACCACACAAGGACACGAG 5047

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5048 TTGGCCATGATGACATACCATGGCAATTCGGCCTTCTGGCAACCAAAAGCAATGAAAC 5107

4558 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTCTGGAGACGATGGT 4617
41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
4618 TATGCCAAGGATGCAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGTGATGGGAG 4677
61 LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProphe 80
4678 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTC 4737
81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
4738 CTCAACACCCGCAAGCACCTGTATGAGCTGTCTTCCAAATTGACCAGGAGCTCTATCTGTTT 4797
101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
4798 GATACCAACCCGCAAGCACCTGTATGAGCTGTCTTCCAAATTGACCAGGAGCTCTATCTGTTT 4857
121 PheThrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
4858 TTCACCTACACTGGGACGGGACATCATCACTCATCACAGACAACAATGGCAACATGGTA 4917
141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGlnVal 160
4918 AATGTCCGCCGAGACTCTACTGGGATGCCCTCTGGTGGTGGTCCCAGATGGCCAGGTG 4977
161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
4978 TACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCACACAGGACACGAG 5037
181 LeuAlaMetMetThrThrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
5038 TTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAAAC 5097
201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
5098 GGATGGACAACATTTTATGAGTACGACAGCTTTGGCCCGCTGACAAATGTGACCTTCCCT 5157
221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
5158 ACTGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAG 5217
241 ThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
5218 ACCTCCAGCAAGGATGATGTACCATATAACCAACCAACCTGTGTGCTCAGGCGCTTCTAC 5277

RESULT 5

US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderne, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E

201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
5108 GGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCTGACAAAATGTGACCTTCCCT 5167
221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
5168 ACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAG 5227
241 ThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
5228 ACCTCCAGCAAGGATGATGTACCATATAACCAACACCTGTCTGCTCAGGCGCTTCTAC 5287
261 Thr 261
5288 ACA 5290

RESULT 4

US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 9.08e-170 Length: 8355
Score: 1379.00 Matches: 260
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.64% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-383-201-55 (1-8355)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
4498 AAGATCAACCCGATCAGGACAGGTACCACTAGTGGAGAGATCTCACTGTTGCTGGGGCC 4557
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40

```

; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:
Pred. No.:      2.49e-111      Length:      8473
Score:          935.50        Matches:     172
Percent Similarity: 79.01%    Conservative: 35
Best Local Similarity: 65.65%  Mismatches:  54
Query Match:      67.59%     Indels:       1
DB:              17         Gaps:         1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-39 (1-8473)

QY      1  Lys1LeAsnArgIleArgGlnValThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db      4317 AAAATTAAACCGGATAAGGCAGGTCAACACAGATGGAGAAATCTCTTAGTGGCCGGAATA 4376

QY      21  ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db      4377 CCTTCAGAGTGTGACTGCAGAAATGATGCCAACTGTGACTGTACCAGAGTGGAGATGGC 4436

QY      41  TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db      4437 TAGCCCAAGGATGCCAAACTCAGTGCCCCCATCTCCCTGGCTGCTTCTCCAGATGGTACA 4496

QY      61  LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db      4497 CTGTATATTGCAGATCTAGGGAATATCCGGATCCGGGCTGTGTCAAAGAATAAGCCTTTA 4556

QY      81  LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db      4557 CTTAACCTATGAACCTTCTATGAAGTTGCGTCTCCAACCTGATCAAGAATCTATACATCTTT 4616

QY      101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db      4617 GACATCAATGGTACTCACCAATATACTGTAAGTTTAGTCACTGGTGATTACCTTTACAAT 4676
```

```

QY      121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db      4677 TTAGCTACAGCAATGACAATGATATTAAGTGTGTGACAGACAGCAATGGCAACCCCTT 4736

QY      141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db      4737 AGAATTAGACGGGACCCCAATCGCATGCCAGTTCGAGTGGTGTCTCTCTGATAACCAAGTG 4796

QY      161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db      4797 ATATGGTTGACAATAGGAACAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAA 4856

QY      181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db      4857 TTAGTTTGTGTTTACCTACCATGGCAATAGTGGCCTTTTAGCCACTAAAGATGATGAAACT 4916

QY      201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db      4917 GGATGGACAACGTTTGTGACTATGACAGTGAAGTCTGTGACAAATGTACGTTTCCA 4976

QY      221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db      4977 ACTGGAGTGTCAAAACCTGTCATGGGGACATGGACAAGGCTATCACAGTGGACATTGAG 5036

QY      241 ThrSerSerLys--AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db      5037 TCATCTAGCCGAGAGAAGAAGATGTCAGCATCACTTCAAATCTGTCTCGATCGATTCTTTC 5096

QY      260 TyrThr 261
Db      5097 TACACC 5102
```

```

RESULT 6
US-10-038-854-41
; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennnda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
```

```
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 2.49e-111 Length: 8487
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.59% Indels: 1
DB: 17 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-41 (1-8487)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4313 AAAATTAAACCGGATAAGGCAGGTCAACACAGATGGAGAAATCTCTTAGTGGCGGAATA 4372

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4373 CCTTCAGAGTGTGACTGCAAAATGATGCCAACTGTGACTGTTACCAGAGTGAGATGGC 4432

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4433 TACGCCAAGGATGCCAAACTCAGTGCCCCATCTCCCTGGCTGCTTCTCCAGATGGTACA 4492

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4493 CTGTATATTGCAGATCTAGGGAATATCCGGATCCGGCTGTGTCAAAGAAATAGCCCTTTA 4552

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4553 CTTAACCTCTATGAACCTTCTATGAAGTTGCGTCTCCAACTGATCAAGAACTCTACATCTT 4612

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4613 GACATCAATGGTACTCACCAATATATACTGTAAAGTTTAGTCTGCTGATTAACCTTTACAAT 4672

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4673 TTAGCTACAGCAATGACAAATGATATTAATCTGCTGTGACAGACAGCAATGGCAACACCTT 4732

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 4733 AGAATTAGACGGGACCCAAATCGCATGCCAGTTCCAGTGGTGTCTCTCTGATAACCAAGTG 4792

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4793 ATATGGTTGACAATAGGAACAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAA 4852

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 4853 TTAGTTTGTACTTACCATGGCAATAGTGGCCTTTTAGCCCACTAAAAGTGAAGAACT 4912

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 4913 GGATGGACAACGTTTTTGAAGTATGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4972

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 4973 ACTGGAGTGGTCACAAACCTGCATGGGACATGGACAAAGGCTATCAGATGGACATGGAG 5032

QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5033 TCATCTAGCCGAGAGAAGATGTCAGCATCACTTCAAATCTGCTCGATCGATTCCTTC 5092

QY 260 TyrThr 261
Db 5093 TACACC 5098

RESULT 7
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Bsha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
```



```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.:      2.57e-111      Length:      8645
Score:          935.50         Matches:    172
Percent Similarity: 79.01%     Conservative: 35
Best Local Similarity: 65.65%  Mismatches:  54
Query Match:    67.59%        Indels:     1
DB:             17            Gaps:        1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-37 (1-8645)

QY      1  LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db      4489 AAAATTAAACCGGATAGGCAGGTCACAACAGATGGAGAAATCTCCTTAGTGGCCGGAATA 4548

QY      21  ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db      4549 CCTTCAGAGTGTGACTGCAAAATGATGCCAACTGTGACTGTATTACCAGAGTGGAGATGGC 4608

QY      41  TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db      4609 TAGCCCAAGGATGCCAAACTCAGTGCCCCATCTCCCTGGCTGCTTCTCCAGATGGTACA 4668

QY      61  LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db      4669 CTGTATATTGCAGATCTAGGGAATATCCGGATATCCGGATCCGGCTGTGTCAAAGAATAAGCCTTTA 4728

QY      81  LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db      4729 CTTAACTCTATGAACCTTCTATGAAGTTGCGTCTCCAAGTATCAAGAACTCTACATCTTT 4788

QY      101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db      4789 GACATCAATGGTACTACCAATATATCTGTAAGTTTAGTCACTGGTGATTACCTTTACAAT 4848

QY      121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db      4849 TTTAGTACAGCAATGACAATGATATTAATCTGCTGTGTGACAGACAGCAATGGCAACCCCTT 4908

QY      141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db      4909 AGAATTAGACGGGACCCAAATCGCATGCCAGTTCGAGTGTGTCTCTGATACCAAGTGTG 4968

QY      161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db      4969 ATATGTTGACAATAGGAACAATGGATGTTTGAAGGCGATGCTGCTCAAGGACTGGAA 5028

QY      181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db      5029 TTAGTTTGTGTTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAGTGATGAACACT 5088

QY      201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db      5089 GGATGGACAACGTTTTTTTGTACTATGACAGTGAAGGTGCTGTCACAAATGTTACGTTTCCA 5148

QY      221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db      5149 ACTGGAGTGGTCAACAACCTGCATGGGGACATGGACAAGGCTATCACAGTGGACATTGAG 5208

QY      241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db      5209 TCATCTAGCCGAGAGAAGATGTCCAGCATCACTTCAATCTGTCTCGATCGATTCTTTC 5268

QY      260 TyrThr 261
Db      |||||
```

```
Db      5269 TACACC 5274

RESULT 8
US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:
Pred. No.:      2.58e-111      Length:      8675
Score:          935.50         Matches:    172
Percent Similarity: 79.01%     Conservative: 35
Best Local Similarity: 65.65%  Mismatches:  54
Query Match:    67.59%        Indels:     1
```

```
DB: 17 Gaps: 1
US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-038-854-35 (1-8675)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4501 AAAATTAAACCGGATAAGGCAGGTCACAACAGATGGAGAAATCTCCTTAGTGGCCGGAATA 4560
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4561 CCTTCAGAGTGTGACTGCAAAAATGATGCCAACTGTGACTGTGTACAGAGTGGAGATGGC 4620
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4621 TACGCCAAGGATGCCAAACTCAGTGCCCATCCTCCCTGGCTGCTTCTCCAGATGGTACA 4680
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4681 CTGTATATTGCAGATCTAGGGAATATCCGGATCCGGGTGTGTCAAGAATAAGCCTTTA 4740
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4741 CTTAACCTCTATGAACCTTCTATGAAGTTGCGTCTCCAACCTGATCAAGAACTTACATCTTT 4800
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4801 GACATCAATGGTACTACCAATATACTGTAAAGTTAGTCACTGGTGATACCTTTTACAAT 4860
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4861 TTTAGCTACAGCAATGACAATGATATTAATCTGCTGTGTGACAGACAGCAATGGCAACACCTTT 4920
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4921 AGAATTAGACGGGACCAAAATCGCATGCCAGTTCGAGTGGTGTCTCCTGATAACCAAGTG 4980
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4981 ATATGGTTGACAATAGGAACAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAA 5040
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5041 TTAGTTTTGTTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAGTGATGAAACT 5100
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5101 GGATGGACAACAGTTTTTTGACTATGACAGTGAAGGTCTGTGACAAATGTTACGTTTCCA 5160
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5161 ACTGGAGTGGTCACAAACCTGCGATGGGACATGGACAAAGGCTATCACAGTGGACATGAG 5220
QY 241 ThrSerSerLys--AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5221 TCATCTAGCCGAGAAGAAGATGTCAGCATCACTTCAAAATCTGTCTCGATCGATTCTTTC 5280
QY 260 TyrThr 261
Db 5281 TACACC 5286
RESULT 9
US-10-295-027-927
; Sequence 927, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevez, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
```

```
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 927.
; LENGTH: 12880
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-927

Alignment Scores:
Pred. No.: 4.16e-106 Length: 12880
Score: 898.00 Matches: 162
Percent Similarity: 78.46% Conservative: 42
Best Local Similarity: 62.31% Mismatches: 56
Query Match: 64.88% Indels: 0
DB: 16 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x US-10-295-027-927 (1-12880)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4406 AAAAGTAAACCGCATTTCAGCAAGTAACCAATGGGAGATCTACATCATCGTGGTGCC 4465
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4466 CCCACTGACTGTGACTGCAAAAATTTGATCCAAACTGTGACTGTTTTTTCAGGTGATGGTGGC 4525
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4526 TATGCCAAAGATGCAAAAGATGAAAGCCCTTCCTCCTTAGCAGTGTGCGCTGATGGAACC 4585
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4586 CTCTATGTGGCAGACCTCGGAAATGTTTCAATTCGTACCATCAGCAGGAACCAAGCCAC 4645
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4646 CTGAATGACATGAACATTTATGAGATTGCTTCCACCCGCTGATCAGGAACCTGACCAGTTC 4705
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4706 ACTGTAAATGGAAACCCACCTACACACCTGAACCTTGATGATAACAGGACTATGTTTATAAC 4765
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4766 TTCACTACAATTTCTGAAGGTGACTTGGGCGGATTACCAAGCAGCAATGGCAATTCAGTG 4825
```

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGlnVal 160
Db : : : : :
4826 CACATTCGCCGTGATGCAGGCGGAATGCCGTATGGCTTGCTGGCGGACAAGTA 4885
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db : : : : :
4886 TACTGGCTGACTATAAGCAGCAATGGAGTCTGAAAAGAGTGTGAGCCCAAGGCTATAAT 4945
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db : : : : :
4946 CCGGCTTAATGACCTATCCAGGAAACACAGGGCTTCTGGCTACCAAAAGTAACGAAAAAT 5005
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db : : : : :
5006 GGATGGACAACCGTTTATGAGTATGACCCCGAGGGACACCTGGAGAGCTGACAAAAGTGGAGCTAGAT 5125
QY 221 ThrGlyGlnValSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db : : : : :
5066 ACTGGAGAGGTGAGCAGCTTCCACAGTGACCTGGAGAGCTGACAAAAGTGGAGCTAGAT 5125
QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db : : : : :
5126 ACTTCCAACCGTGAAAATGTCTCATGTCAACCACTTGACGGCAACTAGTACCATATAT 5185

RESULT 10
US-09-808-602-78
; Sequence 78, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 78
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-808-602-78

Alignment Scores:
Pred. No.: 4.31e-102 Length: 8689
Score: 865.50 Matches: 163
Percent Similarity: 77.10% Conservative: 39
Best Local Similarity: 62.21% Mismatches: 59
Query Match: 62.54% Indels: 1
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-808-602-78 (1-8689)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db : : : : :
4594 AAGATCAACCGCTACGCCAGGTCCACCACGAGAGATCTGCCTCTTAGCCGGGCA 4653
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db : : : : :
4654 GCCTCAGACTGTGACTGC AAAAATGACGTCAACTGCATCTGCTATTCGGGAGATGACGCA 4713
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60

Db 4714 TACGCCACGGATGCCATCTTGAACCTCCCGTCTCTTCTTAGCTGGCTCCGGATGGCACC 4773
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db : : : : :
4774 ATCTACATCGCAGACCTCGGGAATATCCGGATCAGGGCGGTACGAAACAAACCTGTT 4833
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db : : : : :
4834 CTTAACCGGTTCAACCAAGTATGAGGCTGCGTCTCCGGGAGAACAGGAACGTGACGTGTC 4893
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db : : : : :
4894 AACCGCGATGGTATCCATCAGTACACCGTGAGCGCTGGTACCGGGGAGTACTTATACAAT 4953
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db : : : : :
4954 TTCACCTACAGCGCTGACAATGATGTCAACCGAGTTGATTGATGACAACAACGGGAATTCCTTA 5013
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGlnVal 160
Db : : : : :
5014 AAGATCCGCGGACAGCAGTGGCATGCCCGACACCTGCTCATGCTGATATATCAGATC 5073
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db : : : : :
5074 ATCACCTTACGGTGGGACCAACGGAGGCGCTCAAGCGGTGTCAACGCAGAACCTGGAG 5133
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db : : : : :
5134 CTGGGCTCATGACTATGATGGGAACACTGGACTCTCTAGCCACCAAGAGCGATGAAACC 5193
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db : : : : :
5194 GGATGGACAACCTTTTATGACTATGACCACGAGGCGCTGTGACCAATGTGACTCGCCCC 5253
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db : : : : :
5254 ACGGGGTGGTGACAGCCTGCACCGGGAATGGAGAAATCCATCACCGTTGACATTGAG 5313
QY 241 ThrSerSerLysAsp--AspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db : : : : :
5314 AACTCCAACCGTGATAACGATGTCACTGTGATTACCAACCTCTCTTCAGTGGAGGCCTCC 5373
QY 260 TyrThr 261
Db : : : : :
5374 TACACC 5379

RESULT 11
US-09-800-198-66
; Sequence 66, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 66
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-800-198-66

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5171 GGATGACAACTTTTATGACTATGACACGAGGCGGTCTGACCAATGTGACCGGCC 5230
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5231 ACGGCGTGTGACCACTCTGCACCGGAAATGGAGAAATCTATCACCATTGACATTGAG 5290
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5291 AACTCCAACCGGATGATGACGTCACTGTGATCACCAACCTCTCTCCGTGGAGGCCTCC 5350
QY 260 TyrThr 261
Db 5351 TATACA 5356

RESULT 13
US-09-808-602-77
; Sequence 77, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-808-602-77

Alignment Scores:
Pred. No.: 5.95e-102 Length: 8797
Score: 864.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.46% Indels: 1
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-808-602-77 (1-8797)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4571 AAGATCAACCGCTACGCCAAGTCAACCACCAATGGAGAGATCTGCCTCTTAGCCGGGCG 4630
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4631 GCCTCAGACTGTGACTGCAAAACGATGTCAACTGCATCTGTACTCGGAGATGACGCT 4690
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4691 TACGCCACGGACGCCATCCTGAACCTCGCCCTCTCTTAGCCGTGGCTCGGATGGCACC 4750
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProPhe 80
Db 4751 ATCTACATTGACACCTTGGGAATATCCGATCAGGCGGTGAGCAAAATAAACCCGTT 4810
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

Db 4811 CTTAACGCATTCAACCACTATGAGGCTGCATCTCCGGGAGAACAGGAATTGTACGTGTT 4870
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4871 AACGCTGATGGTATCCATCAGTACACTGTGAGTCTGGTGAATGGGAGTACTTGTACAAT 4930
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4931 TTCACATACAGCGCTGACATGACATGACGTACCGAGTTGATTGACAAACACGGGAATCCCTA 4990
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 4991 AAGATCCGCGGACAGACGATGGCATGCCCGCCACCTGCTCATGCCGGATAATCAGATT 5050
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5051 ATCACCCTTACTGTGGGCACCAATGGAGGCCTCAAAGCCGTGTCCACTCAGAACCTGGAG 5110
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5111 CTGGGCCTCATGACTTATGATGGAAACACTGGACTCCTAGCCACCAAGAGTGAATAACC 5170
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5171 GGATGGACAACTTTTATGACTATGACCACGAGGCGCTGTGACCAATGTGACCCGCCCC 5230
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5231 ACGGCGTGTGACCACTGTGACCGGAAATGGAGAAATCTATCACCATTGACATTGAG 5290
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5291 AACTCCAACCGGATGATGACGTCACTGTGATCACCAACCTCTCTCCGTGGAGGCCTCC 5350
QY 260 TyrThr 261
Db 5351 TATACA 5356

RESULT 14
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62

Alignment Scores:
Pred. No.: 5.95e-102 Length: 8797
Score: 864.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.46% Indels: 1
DB: 10 Gaps: 1

```
US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-800-198-62 (1-8797)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4571 AAGATCAACCGCTACGCCAAGTCACCAATGGAGAGATCTGCCTCTTAGCCGGGCG 4630
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4631 GCCTCAGACTGTGACTGCAAAAACGATGTCAACTGTCATCTGCTACTCGGAGATGACGCT 4690
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4691 TAGCCACGGACGCCATCCTGAACCTGCCGTCTCTTAGCCGTGGCTCCGGATGGCACC 4750
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db 4751 ATCTACATTGCAGACCTTGGGAATATCCGGATCAGGGCGGTGACGAAAAATAAACCCGTT 4810
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4811 CTTAACGCATTCAACCATGAGGCTGCATCTCCGGGAGAACAGGAATTGTACGTGTTT 4870
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4871 AACGCTGATGGTATCCATCAGTACACTGTGAGTCTGGTGACTGGGGAGTACTTGTACAAT 4930
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4931 TTCACATACAGCGCTGACCAATGACGTACCGAGTTGATTGACAAACACGGGAATCCCTA 4990
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4991 AAGATCCGCGGACAGACGTGGCATGCCCGCCACCTGCTCATGCCGATAATCAGATT 5050
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5051 ATCACCTTACTGTGGGCACCAATGGAGGCCCTCAAAGCCGTGCCACTCAGAACCTGGAG 5110
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5111 CTGGCCCTCATGACTTATGATGGGAACACTGGACTCCTAGCCACCAAGAGTGATGAAACC 5170
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5171 GGATGGACAACTTTTATGACTATGACCCAGGGCCGCTGACCAATGTGACCCGCCCC 5230
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5231 ACGGGCGTGGTGACCACTCTGCACCGGAAAATGGAGAAATCTATCACCATTGACATTGAG 5290
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5291 AACTCCAACCGGATGATGACGTCACTGTGATCACCAACCTCTCCTCCGTGGAGGCCTCC 5350
QY 260 TyrThr 261
Db 5351 TATACA 5356
```

RESULT 15
US-09-800-198-65
; Sequence 65, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697

```
; CURRENT APPLICATION NUMBER: US/09/800,198  
; CURRENT FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 98  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 65  
; LENGTH: 8797  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-800-198-65  
  
Alignment Scores:  
Pred. No.: 5.95e-102 Length: 8797  
Score: 864.50 Matches: 162  
Percent Similarity: 77.10% Conservative: 40  
Best Local Similarity: 61.83% Mismatches: 59  
Query Match: 62.46% Indels: 1  
DB: Gaps: 1
```

```
US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-800-198-65 (1-8797)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4571 AAGATCAACCGCTACGCCAAGTCACCAATGGAGAGATCTGCCTCTTAGCCGGGCG 4630
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4631 GCCTCAGACTGTGACTGCAAAAACGATGTCAACTGTCATCTGCTACTCGGAGATGACGCT 4690
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4691 TAGCCACGGACGCCATCCTGAACCTGCCGTCTCTTAGCCGTGGCTCCGGATGGCACC 4750
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db 4751 ATCTACATTGCAGACCTTGGGAATATCCGGATCAGGGCGGTGACGAAAAATAAACCCGTT 4810
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4811 CTTAACGCATTCAACCATGATGAGGCTGCATCTCCGGGAGAACAGGAATTGTACGTGTTT 4870
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4871 AACGCTGATGGTATCCATCAGTACACTGTGAGTCTGGTGACTGGGGAGTACTTGTACAAT 4930
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4931 TTCACATACAGCGCTGACCAATGACGTACCGAGTTGATTGACAAACACGGGAATTCCTTA 4990
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4991 AAGATCCGCGGACAGACGTGGCATGCCCGCCACCTGCTCATGCCGATAATCAGATT 5050
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5051 ATCACCTTACTGTGGGCACCAATGGAGGCCCTCAAAGCCGTGCCACTCAGAACCTGGAG 5110
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5111 CTGGCCCTCATGACTTATGATGGGAACACTGGACTCCTAGCCACCAAGAGTGATGAAACC 5170
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5171 GGATGGACAACTTTTATGACTATGACCCAGGGCCGCTGACCAATGTGACCCGCCCC 5230
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5231 ACGGGCGTGGTGACCACTCTGCACCGGAAAATGGAGAAATCTATCACCATTGACATTGAG 5290
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5291 AACTCCAACCGGATGATGACGTCACTGTGATCACCAACCTCTCCTCCGTGGAGGCCTCC 5350
```


Qy 260 TyrThr 261
|||
Db 5351 TATACA 5356

Search completed: August 14, 2004, 19:19:36
Job time : 604.097 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 77.2165 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750
Perfect score: 1384
Sequence: 1 KINRIRQVTTSGEISLVAGA.....SSKDDVTITNLSASGAFYT 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1.1 258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match %	Length DB ID	Description	
1	124.5	9.0	4881	4	US-09-252-991A-7234
2	115	8.3	31960	4	US-09-453-702B-11
3	112.5	8.1	2415	4	US-09-328-352-1419
4	105.5	7.6	4761	4	US-09-543-681A-1157
5	105	7.6	640681	4	US-09-790-988-1
6	101.5	7.3	885	4	US-09-446-301A-3
7	101.5	7.3	885	4	US-09-099-932-3
8	101.5	7.3	1607	4	US-09-446-301A-16
9	101.5	7.3	1607	4	US-09-099-932-50
10	98.5	7.1	1515	4	US-09-252-991A-11708
11	98	7.1	4860	4	US-09-328-352-3221
C 12	97.5	7.0	1260	4	US-09-252-991A-15163

C 13	97.5	7.0	2181	4	US-09-252-991A-14909	Sequence 14909, A
14	97.5	7.0	3387	4	US-09-252-991A-14461	Sequence 14461, A
C 15	96.5	7.0	2236	4	US-09-221-017B-482	Sequence 482, Appl
16	96	6.9	3015	4	US-09-206-942-56	Sequence 56, Appl
17	96	6.9	3033	4	US-09-206-942-54	Sequence 54, Appl
C 18	94.5	6.8	11613	4	US-09-453-702B-42	Sequence 42, Appl
C 19	94	6.8	642	4	US-09-252-991A-7031	Sequence 7031, Ap
C 20	94	6.8	684	4	US-09-252-991A-11528	Sequence 11528, A
21	94	6.8	1593	4	US-09-252-991A-7095	Sequence 7095, Ap
22	94	6.8	2715	4	US-09-543-681A-2771	Sequence 2771, Ap
C 23	94	6.8	4453	4	US-09-453-702B-90	Sequence 90, Appl
24	93.5	6.8	795	4	US-09-134-001C-2198	Sequence 2198, Ap
25	93.5	6.8	804	4	US-09-134-001C-1702	Sequence 1702, Ap
C 26	93	6.7	3900	4	US-09-023-655-1420	Sequence 1420, Ap
C 27	93	6.7	11935	4	US-09-634-238-401	Sequence 401, App
28	92.5	6.7	1785	4	US-09-489-039A-6762	Sequence 6762, Ap
29	92	6.6	5319	1	US-08-169-927-1	Sequence 1, Appli
30	92	6.6	25165	4	US-09-453-702B-39	Sequence 39, Appl
31	91.5	6.6	2781	4	US-09-313-677-1	Sequence 1, Appli
32	91.5	6.6	2799	4	US-09-313-677-18	Sequence 18, Appl
33	91.5	6.6	7026	4	US-09-313-677-20	Sequence 20, Appl
34	91.5	6.6	7344	4	US-09-313-677-16	Sequence 16, Appl
35	91	6.6	49617	4	US-09-596-002-28	Sequence 28, Appl
36	90.5	6.5	5738	1	US-08-409-995-3	Sequence 3, Appli
37	90.5	6.5	5738	3	US-08-685-467-3	Sequence 3, Appli
38	90.5	6.5	7291	3	US-08-913-942-3	Sequence 3, Appli
39	89.5	6.5	11679	4	US-09-328-352-1377	Sequence 1377, Ap
40	89	6.4	1443	4	US-09-724-623-15	Sequence 15, Appl
C 41	89	6.4	1683	4	US-09-252-991A-6977	Sequence 6977, Ap
C 42	89	6.4	1683	4	US-09-252-991A-11588	Sequence 11588, A
43	89	6.4	2925	4	US-09-252-991A-7069	Sequence 7069, Ap
44	89	6.4	2997	4	US-09-252-991A-11853	Sequence 11853, A
45	88.5	6.4	2745	4	US-09-817-514A-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-252-991A-7234
; Sequence 7234, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7234
; LENGTH: 4881
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7234

Alignment Scores:				
Pred. No.:	6.01e-05	Length:	4881	
Score:	124.50	Matches:	53	
Percent Similarity:	41.33%	Conservative:	40	
Best Local Similarity:	23.56%	Mismatches:	98	
Query Match:	9.00%	Indels:	35	
DB:	4	Gaps:	7	
US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-7234 (1-4881)				
QY	51	SerSerLeuAlaValCysAlaAspGly-----GluLeuTyrValAlaAspLeuGlyAsn	68	


```

RESULT 3
US-09-328-352-1419
; Sequence 1419, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1419
; LENGTH: 2415
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1419

Alignment Scores:
Pred. No.:          0.000664          Length:          2415
Score:              112.50            Matches:          47
Percent Similarity: 36.11%            Conservative:     18
Best Local Similarity: 26.11%          Mismatches:      81
Query Match:        8.13%              Indels:          35
DB:                  4                Gaps:            5

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-328-352-1419 (1-2415)
QY 91 SerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLeuTyrThrGln 110
Db 767 AGTCCTGTGGAGAAAACACTGGTG-----CAATTGTACAAACA 805
QY 111 SerLeuProThrGlyAspTyrLeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThr 130
Db 806 CAGATAGCTTT-GGTACAAACAGAACTAATCTATTATTAGCAGCAGGAGATTTTAGTGCTGAT 864
QY 131 LeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSerThrGlyMetPro 150
Db 865 TTAAAAACA-----CTGGTTGCACGTAAAGGTACA----- 894
QY 151 LeuTrpLeuValValProAspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAla 170
Db 895 -----GATGGATCAGTCCTTCAGACAACAACACTGAATTGAAAACGCA 936
QY 171 Leu-----LysSerValThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGly 188
Db 937 CTTTTAAACCAAAAGTTGCTGTATCGGGTAGTTCTACACAAAGTTATACATATAACTCT 996
QY 189 AsnSerGlyLeuLeuAlaThrLysSerAsnGluAsnGlyTrpThrPheTyrGluTyr 208
Db 997 AATACAACTTTAGCAACAGAAAAGATAATAGTATTAGTGGGCAAAAGACATTTAAATAT 1056
QY 209 AspSerPheGlyArgLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArg 228
Db 1057 GATGATACTGGCCGAATTACTTCAATTACCCATCCAGATAGTTCCAGTAGAAACGATCAA 1116
QY 229 -----SerAspThrAspSerSerValHisValGlnVal 239
Db 1117 TATTTCAATTAAAGGACCTCATCGCGTCCCGGACATGGCGAGAAGTTGAGACAACTTAT 1176
QY 240 GluThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 1177 AGTTATTCTTCTTGCTGGCGGTTTTAAAAACAACGACTAATGCAATATTAGTGAAGCAATT 1236

RESULT 4
US-09-543-681A-1157
; Sequence 1157, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
```

```

; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 1157
; LENGTH: 4761
; TYPE: DNA
; ORGANISM: Proteus mirabilis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (4726)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-543-681A-1157

Alignment Scores:
Pred. No.:          0.0143          Length:          4761
Score:              105.50            Matches:          44
Percent Similarity: 36.00%            Conservative:     28
Best Local Similarity: 22.00%          Mismatches:      73
Query Match:        7.62%              Indels:          55
DB:                  4                Gaps:            7

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-543-681A-1157 (1-4761)
QY 85 AsnMetTyrGluLeuSerSerProIleAspGlnGlu---LeuTyrLeuPheAspThrThr 103
Db 2125 AATTTAACCAATGGACCGATCCGGAACAACAATCGACTTGCTATGATGGGACAAAGAT 2184
QY 104 GlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsnPheThrTyr 123
Db 2185 TTTGCGTTTCCCACGGCACAAACCTTGCCAAATGGG---GCGGCTTGGCATGGGAGTAC 2241
QY 124 ThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnVal--- 142
Db 2242 AATGAGCACGGCGATATTCTGTCGTGTGATGATGACCCGCTTGGCCATATCACGGCCTTGGCG 2301
QY 142 ----- 142
Db 2302 TGGGATGACCAAGGCCTGTGTCTGGGCAAGTGATGCTAAGGGTAATGAAACCCACTAT 2361
QY 143 -----ArgArgAspSerThrGlyMetProLeuTrp 152
Db 2362 CGCTATAATGCCCGGTCAGTTAATCGAGCAACGGGACTGTTCCGGTTATCCCACCACC 2421
QY 153 LeuValValProAspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAlaLeuLys 172
Db 2422 TTGACCTACGATGAT-----TGG-----GGACAACTTCGC 2451
QY 173 SerValThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeu 192
Db 2452 TCGTTGACCAATGCACAGAAATGAAACACGACTTACACCTTT---AGTGAAGCGGGTTG 2508
QY 193 LeuAlaThrLysSerAsnGluAsnGlyTrpThrThrPheTyrGluTyrAspSerPheGly 212
Db 2509 TTGTTAACAGAGTGCTTACC GGATGGGACAGAAAACCGTTATGACTATGATGCCACCGGG 2568
QY 213 ArgLeuThrAsnValThrPhePro-----Thr 221
Db 2569 CAATTAGTGGGAATAACGGATGCCGAGAGCGCCATATTCTTGTTCGCCCGTAACCGTCGT 2628
QY 222 GlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGluThr 241
Db 2629 GGGCAAGTGATAGCCCCGACCGCATCCGGCAGGGCATTCGGTTGCATTTTCACTATGATACT 2688

RESULT 5
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEKI
; APPLICANT: HATTORI, MASAHIRA
```

; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; PRIOR FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Alignment Scores:
Pred. No.: 30.5 Length: 640681
Score: 105.00 Matches: 44
Percent Similarity: 41.78% Conservative: 45
Best Local Similarity: 20.66% Mismatches: 88
Query Match: 7.59% Indels: 36
DB: 4 Gaps: 9

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-790-988-1 (1-640681)

QY 27 LysAsnAspAlaAsnCysAspCysPheSerGlyAspGlyTyrAlaLysAspAlaLys 46
Db 219524 AAAGAAGATTTCAGATATTTTGTGGAACACGAGAGATTTTCTATTGATGAAAAATCAAAA 219583
QY 47 LeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGluLeuTyrValAlaAspLeu 66
Db 219584 -----CAATCTATTAAACAGAAAGA 219604
QY 67 GlyAsnIleArgIleArgPheIleArgLysAsnLysPropheLeuAsnThr--GlnAsn 85
Db 219605 GGATTGATTAAAGTTCAGAAAAATATTATTGATAAAAAATTAATGAATACAGGTGAATCA 219664
QY 86 MetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLys 105
Db 219665 TTATATTCTTCAATAATAATAATAATAATGATCATCATGATTATCTGCACCTCGTGGCAT 219724
QY 106 HisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsnPheThrTyrThrGly 125
Db 219725 AAATTATTGTTCGAAATGTT-----GATTATCTTGTAATA----- 219760
QY 126 AspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAsp 145
Db 219761 GATAATAGTGTGATTATTGTTGATGAACATACAGGTGCGCACTATGCCAGGAAGAATGG 219820
QY 146 SerThrGlyMetProLeuTrpLeuValProAspGlyGlnValTyrTrpValThrMet 165
Db 219821 TCAGATGGATTACATCAAGCAATAGAGCAAGAAAGAAAT-----GTTCTATA 219868
QY 166 ---GlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMet 184
Db 219869 AAAAATGAAAATCAAACTTTAGCATCTATTACTTTTCAAAATTAT-----TTTCGT 219919
QY 185 ThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsnGlyTrpThrThr 204
Db 219920 TTATATGAAAAAATAGCAGGTATGACGGTACTGCTGAAACTGAATCG----- 219967
QY 205 PheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPheProThrGlyGlnVal 224
Db 219968 ---TTTGAGTTTAACTCTATTATATAATCTTGATACTATTGTAATACCAACAATAGA--- 220021
QY 225 SerSerPheArgSerAspThrAspSerSerValHisVal 237
Db 220022 AAAATGATACGTAAAGATTTCAGACTTAGTATATATG 220060

RESULT 6
US-09-446-301A-3
; Sequence 3, Application US/09446301A
; Patent No. 6506893

; GENERAL INFORMATION:
; APPLICANT: EL SOLH, NEVINE
; APPLICANT: ALLIGNET, JEANINE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: 03715-0059
; CURRENT APPLICATION NUMBER: US/09/446,301A
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Staphylococcus sp.
US-09-446-301A-3

Alignment Scores:
Pred. No.: 0.00346 Length: 885
Score: 101.50 Matches: 49
Percent Similarity: 41.45% Conservative: 31
Best Local Similarity: 25.39% Mismatches: 83
Query Match: 7.33% Indels: 30
DB: 4 Gaps: 10

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-446-301A-3 (1-885)

QY 68 AsnIleArgIleArgPheIleArgLysAsnLysPropheLeuAsnThr-GlnAsnMetTy 87
Db 48 TCCATACGGTATAACTTCATCAGAAAGCGAAAGGTATGGTTCACACACATAGGCAAA 107
QY 87 rGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLe 107
Db 108 CAAAATCAGCAGT-----CTAGATCAGATGGTAGG---AT 140
QY 107 uTyrThrGlnSerLeuProThrGlyAsp---TyrLeuTyrAsnPheThrTyrThrGlyAs 126
Db 141 AAAAGAATTTCGAAAGTCTCTACCCCTGATGCTAAAGTGATGTGTTTAAATTGTATCTCACT 200
QY 126 pGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSe 146
Db 201 TGGAGACATA---TGGTTTACAGAGAATGGTGCAATAATAAATCGGAAAGCTCTCAAAAAA 257
QY 146 rThrGlyMetProLeuTrp---LeuValValProAspGlyGlnValTyrTrpValThrMe 165
Db 258 AGTGGCTTTACAGATATCCATTGCCACAGCCGAGTCTGGTCTTACGGAATAACGGA 317
QY 165 tGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMetTh 185
Db 318 AGGTCTAAATGGCGATATATGTTTACCAATTGAATGGAGATCGTATAGGAAAGTTGAC 377
QY 185 rTyrHisGly-----AsnSerGly-----LeuLeuAl 194
Db 378 AGCTGATGGGACTATTATGAATATGATTTGCCAAATAAGGGATCTTATCCTGCTTTTAT 437
QY 194 aThrLysSerAsnGluAsnGly---TrpThrThrPheTyrGluTyrAspSerPheGlyAr 213
Db 438 TACTTTAGGTTTCGATAACGCACTTTGGTTTCACGGAGAACCAAAATAATTCTATTGGAAG 497
QY 213 gLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThrAspSe 233
Db 498 GATTACAAAT-----ACAGGGAATTAGAGAATAATCTCTACCAACAATATGC 545
QY 233 rSerValHisValGlnValGluThrSerSerLysasp 245
Db 546 AGCGGCTCCAGTGGGTATC---ACTAGTGGTAACGAT 579

RESULT 7
US-09-099-932-3
; Sequence 3, Application US/09099932
; Patent No. 6570001
; GENERAL INFORMATION:
; APPLICANT: El Solh, Nevine


```

; APPLICANT: Allignet, Jeanine
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: 03495.0173-00000
; CURRENT APPLICATION NUMBER: US/09/099,932
; CURRENT FILING DATE: 1998-06-19
; EARLIER APPLICATION NUMBER: 60/050,380
; EARLIER FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 885
; TYPE: DNA
; ORGANISM: Staphylococcus
US-09-099-932-3
```

```

Alignment Scores:
Pred. No.:      0.00346      Length:      885
Score:          101.50      Matches:      49
Percent Similarity: 41.45%      Conservative: 31
Best Local Similarity: 25.39%      Mismatches: 83
Query Match:      7.33%      Indels:      30
DB:              4      Gaps:      10
```

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-099-932-3 (1-885)

```

QY 68 AsnIleArgIleArgPheIleArgLysAsnLysProPheLeuAsnThr-GlnAsnMetTy 87
Db 48 TCCATACGGTATAACTTCATCAGAGACGGAAGGTATGGTTCACACACATAAGGCAA 107
QY 87 rGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLe 107
Db 108 CAAATCAGCAGT-----CTAGATCAGAGTGGTAGG--AT 140
QY 107 uTyrThrGlnSerLeuProThrGlyAsp---TyrLeuTyrAsnPheThrTyrThrGlyAs 126
Db 141 AAAAGAATTCGAAGTTCCTACCCCTGATGCTAAAGTGATGTGTTTAATTGTATCTTCACT 200
QY 126 pGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSe 146
Db 201 TGGAGACATA---TGGTTTACAGAGAAATGGTGCAAAATAAAATCGGAAGCTCTCAAAAAA 257
QY 146 rThrGlyMetProLeuTrp---LeuValValProAspGlyGlnValTyrTrpValThrMe 165
Db 258 AGGTGGCTTTACAGAAATATCCATTGCCACAGCCGGATTCTGGTCCTTACGGAATAACGGA 317
QY 165 tGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMetTh 185
Db 318 AGGTCTAAATGGCGATATATGGTTTACCCAATTGAATGGAGATCGTATAGGAAAGTTGAC 377
QY 185 rTyrHisGly-----AsnSerGly-----LeuLeuAl 194
Db 378 AGCTGATGGGACTATTATGAATATGATTTGCCAAATAAGGGATCTTATCCTGCTTTTAT 437
QY 194 aThrLysSerAsnGluAsnGly---TrpThrThrPheTyrGluTyrAspSerPheGlyAr 213
Db 438 TACTTTAGGTTCCGATAACGCACCTTGGTTCACGGAGAACCAAAATAATTCTATTGGAAG 497
QY 213 gLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThrAspSe 233
Db 498 GATTACAAAT-----ACAGGGAATATTAGAAGAATATCCTCTACCAACAAATGC 545
QY 233 rSerValHisValGlnValGluThrSerSerLysAsp 245
Db 546 AGCGGCTCCAGTGGGTATC---ACTAGTGGTAACGAT 579
```

RESULT 8

```

US-09-446-301A-16
; Sequence 16, Application US/09446301A
; Patent No. 6506893
; GENERAL INFORMATION:
; APPLICANT: EL SOLH, NEVINE
```

```

; APPLICANT: ALLIGNET, JEANINE
; TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
; TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: 03715-0059
; CURRENT APPLICATION NUMBER: US/09/446,301A
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 1607
; TYPE: DNA
; ORGANISM: Staphylococcus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (39..923, 947..1582)
US-09-446-301A-16
```

```

Alignment Scores:
Pred. No.:      0.00863      Length:      1607
Score:          101.50      Matches:      49
Percent Similarity: 41.45%      Conservative: 31
Best Local Similarity: 25.39%      Mismatches: 83
Query Match:      7.33%      Indels:      30
DB:              4      Gaps:      10
```

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-446-301A-16 (1-1607)

```

QY 68 AsnIleArgIleArgPheIleArgLysAsnLysProPheLeuAsnThr-GlnAsnMetTy 87
Db 86 TCCATACGGTATAACTTCATCAGAGACGGAAGGTATGGTTCACACACATAAGGCAA 145
QY 87 rGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLe 107
Db 146 CAAATCAGCAGT-----CTAGATCAGAGTGGTAGG--AT 178
QY 107 uTyrThrGlnSerLeuProThrGlyAsp---TyrLeuTyrAsnPheThrTyrThrGlyAs 126
Db 179 AAAAGAATTCGAAGTTCCTACCCCTGATGCTAAAGTGATGTGTTTAATTGTATCTTCACT 238
QY 126 pGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSe 146
Db 239 TGGAGACATA---TGGTTTACAGAGAAATGGTGCAAAATAAAATCGGAAGCTCTCAAAAAA 295
QY 146 rThrGlyMetProLeuTrp---LeuValValProAspGlyGlnValTyrTrpValThrMe 165
Db 296 AGGTGGCTTTACAGAAATATCCATTGCCACAGCCGGATTCTGGTCCTTACGGAATAACGGA 355
QY 165 tGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGluLeuAlaMetMetTh 185
Db 356 AGGTCTAAATGGCGATATATGGTTTACCCAATTGAATGGAGATCGTATAGGAAAGTTGAC 415
QY 185 rTyrHisGly-----AsnSerGly-----LeuLeuAl 194
Db 416 AGCTGATGGGACTATTATGAATATGATTTGCCAAATAAGGGATCTTATCCTGCTTTTAT 475
QY 194 aThrLysSerAsnGluAsnGly---TrpThrThrPheTyrGluTyrAspSerPheGlyAr 213
Db 476 TACTTTAGGTTCCGATAACGCACCTTGGTTCACGGAGAACCAAAATAATTCTATTGGAAG 535
QY 213 gLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThrAspSe 233
Db 536 GATTACAAAT-----ACAGGGAATATTAGAAGAATATCCTCTACCAACAAATGC 583
QY 233 rSerValHisValGlnValGluThrSerSerLysAsp 245
Db 584 AGCGGCTCCAGTGGGTATC---ACTAGTGGTAACGAT 617
```

RESULT 9

```

US-09-099-932-50
; Sequence 50, Application US/09099932
; Patent No. 6570001
; GENERAL INFORMATION:
```

APPLICANT: El Solh, Nevine
APPLICANT: Allignet, Jeanine
TITLE OF INVENTION: POLYNUCLEOTIDES AND THEIR USE FOR DETECTING RESISTANCE
TITLE OF INVENTION: TO STREPTOGRAMIN A OR TO STREPTOGRAMIN B AND RELATED
TITLE OF INVENTION: COMPOUNDS
FILE REFERENCE: 03495.0173-00000
CURRENT APPLICATION NUMBER: US/09/099,932
CURRENT FILING DATE: 1998-06-19
EARLIER APPLICATION NUMBER: 60/050,380
EARLIER FILING DATE: 1997-06-20
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 50
LENGTH: 1607
TYPE: DNA
ORGANISM: Staphylococcus
FEATURE:
NAME/KEY: CDS
LOCATION: (39)..(923)
FEATURE:
NAME/KEY: CDS
LOCATION: (947)..(1582)
US-09-099-932-50

Alignment Scores:
Pred. No.: 0.00863 Length: 1607
Score: 101.50 Matches: 49
Percent Similarity: 41.45% Conservative: 31
Best Local Similarity: 25.39% Mismatches: 83
Query Match: 7.33% Indels: 30
DB: 4 Gaps: 10

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-099-932-50 (1-1607)

QY 68 AsnIleArgIleArgPheIleArgLysAsnLysPropheLeuAsnThr-GlnAsnMetTy 87
Db 86 TCCATACGGTATAACTTCATCAGAGACGGAAAGGTATGTTTCACACACATAAGGCAAA 145
QY 87 rGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLe 107
Db 146 CAAATCAGCAGT-----CTAGATCAGAGTGGTAGG---AT 178
QY 107 uTyrThrGlnSerLeuProThrGlyAsp---TyrLeuTyrAsnPheThrTyrThrGlyAs 126
Db 179 AAAAGAATTTCGAAGTTCCTACCCCTGATGCTAAAGTGATGTGTTAATTGATCTTCACT 238
QY 126 pGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSe 146
Db 239 TGGAGACATA---TGGTTTACAGAGAATGGTGCATAATAAAATCGGAAAGTCTCAAAAA 295
QY 146 rThrGlyMetProLeuTrp---LeuValValProAspGlyGlnValTyrTrpValThrMe 165
Db 296 AGGTGGCTTTACAGAATATCCATTGCCACAGCCGGATTCTGGTCTCTTACGGAATAACGGA 355
QY 165 tGlyThrAsnSerAlaLeuLysSerValThrGlnGlyHisGluLeuAlaMetMetTh 185
Db 356 AGGTCTAAATGGCGATATATGTTTACCCTAATTGAATGGAGATCGTATAGGAAAGTTGAC 415
QY 185 rTyrHisGly-----AsnSerGly-----LeuLeuAl 194
Db 416 AGCTGATGGGACTATTATGAATATGATTTGCCAAATAAGGGATCTTATCCTGCTTTAT 475
QY 194 aThrLysSerAsnGluAsnGly---TrpThrThrPheTyrGluTyrAspSerPheGlyAr 213
Db 476 TACTTTAGGTTCCGATAACGCACCTTTGGTTACGGGAGAACCAAAATAATTCTATTGGAAG 535
QY 213 gLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThrAspSe 233
Db 536 GATTACAAAT-----ACAGGGAATTAGAAGAATATCCTCTACCAACAAATGC 583
QY 233 rSerValHisValGlnValGluThrSerSerLysAsp 245
Db 584 AGCGGCTCCAGTGGGTATC---ACTAGTGGTAACGAT 617

RESULT 10

US-09-252-991A-11708
Sequence 11708, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 11708

LENGTH: 1515

TYPE: DNA

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11708

Alignment Scores:

Pred. No.: 0.0188 Length: 1515
Score: 98.50 Matches: 58
Percent Similarity: 37.20% Conservative: 35
Best Local Similarity: 23.20% Mismatches: 104
Query Match: 7.12% Indels: 53
DB: 4 Gaps: 12

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-11708 (1-1515)

QY 10 ThrSerGlyGluIleSerLeuValAlaGlyAlaProSerGlyCysAspCysLysAsnAsp 29
Db 274 ACCTCGGGAGCTACCTTGATCGGCAACAGCTGCCACTAGGGGCTGAGCTTAATAACGAG 333
QY 30 --AlaAsnCysAspCysPheSerGly-----AspAspGlyTyrAlaLysAspAla 45
Db 334 ACAGGAAGCTCGGATTGCCGCCCGCTTACGAAATGAATGACGCTGGCTGTAAGCTTCCA 393
QY 46 LysLeuAsnThrProSerSerLeuAlaVal-----CysAlaAspGlyGluLeu 61
Db 394 GATAAAATGGCCCCGACAAAGGGGCTCCGCCCGCGAGGGTTGCGCAGGAAACCCGGTC 453
QY 62 TyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysPropheLeu 81
Db 454 AACATCACCAAT---GGCAAC-----471
QY 82 AsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeu----- 99
Db 472 AAATATCAGGTCGAGCATGACTTAATAACCCCGATTCCGCTTGCCCGACACTACAATGGC 531
QY 100 PheAspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyr 119
Db 532 CTGGACGGATTATGGGCTCATTCGTTCTCTGCGGCATAACACGCAAGGATGATAGCTAC 591
QY 120 AsnPheThrTyrThrGlyAspGly-----AspIle 129
Db 592 ---CTACTTTACCGGAAGACGGAAAGGTTTCGGAGTTCAGTGGGCGCGGAGAGATCTA 648
QY 130 ThrLeuIleThrAsp-----AsnAsn 136
Db 649 ACTTCCCTACGGGATCTGGGAAAGCTTAGCCGACTAGCAGGAAGATTTTCTTACCTCG 708
QY 137 GlyAsnMetValAsnValArgArgAspSerThrGlyMetProLeuTrpLeuValPro 156
Db 709 GAGCTCAACGAAACTATAGAGTTTGACCCCGTATGGAATAACTCGCCAGGCTCAAGACCAAG 768
QY 157 AspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThr 176
Db 769 GAGGTCGAAAGTATCGTGTAGAGCGGCGGCTAATCTTACGATTAGCGAC-----GAA 822

QY 177 GlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLys 196
Db 823 CGTGGCAACAACCTC---GTCTCTCGGAGGAGCCCAACCATCAACTCCTT-----CGC 873

QY 197 SerAsnGluAsnGlyTyrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsn 216
Db 874 GCTCAGATAGGGGTATATCCATTGAGTACACATACGACAAGGAACAGCGACTCACTTCG 933

QY 217 ValThrPheProThrGlyGlnValSerSer 226
Db 934 GTAACACGAACAGATGGGCAATACAACACG 963

RESULT 11
US-09-328-352-3221
; Sequence 3221, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 3221
; LENGTH: 4860
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-3221

Alignment Scores:
Pred. No.: 0.13 Length: 4860
Score: 98.00 Matches: 64
Percent Similarity: 34.39% Conservative: 44
Best Local Similarity: 20.38% Mismatches: 110
Query Match: 7.08% Indels: 96
DB: 4 Gaps: 15

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-328-352-3221 (1-4860)

QY 7 GlnValThrThrSerGlyGluLeuSerLeuValAlaGlyAlaProSerGlyCysAspCys 26
Db 2566 CAAGGTACAACTGAAGCAAATGCCACAATAATTGTTAAAAATACTGCAGGT---GATGTA 2622

QY 27 LysAsnAspAlaAsnCysAspCysPheSerGly----- 37

Db 2623 ATCGGTTCCGCAATAGCTGATGCA--TCCGGGTATTACAATGTAATTTGAATACGGTT 2679

QY 38 ---AspAspGly-----TyrAlaLysAspAlaLys----- 46
Db 2680 TATGAAGATGGTGAACCTATTAAAAGTAAATTCGCGGATGCTAAAGTAATGAAAGTTCA 2739

QY 47 -----LeuAsnThrProSerSerLeuAlaValCys----- 56
Db 2740 ATTAATATAAATACACCTGATATTACAGCACCTATATTAGCTAAATTTGTTTAATTATGAT 2799

QY 57 -----AlaAspGlyGluLeuTyrValAlaAasp----- 65
Db 2800 GTTAGTACAGATAAAAATTATTTTAAATGCGCCAAGTGATAGTTATATTGTTGAGCAAAA 2859

QY 66 LeuGlyAsnIleArgIleArgPhe---IleArgLysAsnLysProPheLeuAsnThrGln 84
Db 2860 ATTGGTGATGCTTGGGTTTCAGGTAATGTTGAGGAAAAATTCGATGGTTGTTGAATACAGAA 2919

QY 85 -----AsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPheAasp 101
Db 2920 TTCAGAGTAACCTGCTAAAGATCTAGCTGGGAATAGCTCTCAACCACTAACCATTATAATT 2979

QY 102 ThrThrGlyLysHisLeuTyrThrGlnSerLeuProThr----- 114
Db 2980 AATACTGCATCGGGAACCTTATAAGCCAACCTGATCCTACATTTATTCAAATTTATAAGGA 3039

QY 115 -----GlyAspTyrLeuTyr----- 119

Db 3040 TCTATAGGAAATGATTATCTTTATGGTGGTAATGGGATGATACATTAGTTTCTAATACA 3099

QY 120 -----AsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGly 137
Db 3100 GGCTCTGATTATTTGATGGTGGCTCAGGTAATGATACGTTGATTATGGTGGTAATCC 3159

QY 138 AsnMetValAsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAsp 157
Db 3160 AATGTGTATACGGCTTTACAAGGTCAAGCTGGGAATGATATTATATTGTAGAT----- 3213

QY 158 GlyGlnValTyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGln 177
Db 3214 -----AAAGCATTACTTACTTCA 3231

QY 178 GlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSer 197
Db 3232 TCAAGCTCTATTTCATATTTTAGATAAATGCAGCTGAAGAAAATATCCTTCAATTTAAATCA 3291

QY 198 AsnGluAsnGlyTrpThrPheTyrGluTyrAspSerPhe----- 211
Db 3292 GTGTCCTCTGGTGATATTTCACCTTAAGCAGTCGGATTCACTCATCATATAATCATTTCAAT 3351

QY 212 GlyArgLeuThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAspThr 231
Db 3352 GATTGAGCATCGACAATACGTTTTGGTGAGGGCAATTATCTTCTATTGTATTGTAT--- 3408

QY 232 AspSerSerVal-----HisValGlnValGluThrSerSer 243
Db 3409 GATGGAACAGICTGGAATAAAGCACAAATTTGAAGCTAATACT 3450

RESULT 12
US-09-252-991A-15163/c
; Sequence 15163, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15163
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15163

Alignment Scores:
Pred. No.: 0.019 Length: 1260
Score: 97.50 Matches: 30
Percent Similarity: 42.24% Conservative: 19
Best Local Similarity: 25.86% Mismatches: 60
Query Match: 7.04% Indels: 7
DB: 4 Gaps: 3

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-15163 (1-1260)

QY 118 LeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGly 137
Db 856 GTCTGGCGTTACGACTACGACAACGAAGGACGTTTGGTCGCCAGCAGCAGCCCGCTGGGA 797

QY 138 AsnMetValAsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAsp 157
Db 796 CAACTGACCCGCGCGCTACGACCCGCTGGGGCAACTGATCGGCTGGAATTGGCCGAC 737

QY 158 GlyGlnValTyr-----TrpValThrMetGlyThrAsnSerAlaLeuLysSerVal 174
Db 158 GlyGlnValTyr-----TrpValThrMetGlyThrAsnSerAlaLeuLysSerVal 174


```
Db      736 GGCAGCGCCCTGAGCTATGAATACGACGCCCTTGGCGGCAGACCCGTATC-----GCC 683
QY      175 ThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAla 194
Db      682 GATGCCGAGGGCACGCCACGCTGTTTCCTGGGGCAGCGCGAT-----CTGCTGGCG 629
QY      195 ThrLysSerAsnGluAsnGlyTrpThrPheTyrGluTyrAspSerPheGlyArgLeu 214
Db      628 CGGGTTTCGACGCCGCGCGCGGCGGAATTGCTCTACCTGCACGACGAAGCCGCGAGGCTG 569
QY      215 ThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAsp 230
Db      568 GTCGCCCTGACCAACGAGAACGGCGTCCAGGCGCAGTTCGCTACGAC 521

RESULT 13
US-09-252-991A-14909/c
; Sequence 14909, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14909
; LENGTH: 2181
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14909

Alignment Scores:
Pred. No.:          0.044          Length:          2181
Score:             97.50          Matches:         30
Percent Similarity: 42.24%        Conservative:   19
Best Local Similarity: 25.86%     Mismatches:    60
Query Match:       7.04%          Indels:         7
DB:                  4            Gaps:           3

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-14909 (1-2181)
QY      118 LeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnGly 137
Db      678 GTCTGGCGTTACGACTACGACAAACGAGGACGTTTGGTCGCCAGCAGCCCGCTGGGA 619
QY      138 AsnMetValAsnValArgAspSerThrGlyMetProLeuTrpLeuValValProAsp 157
Db      618 CAACTGACCCGCGCTACGACCCGCTGGGGCAACTGATCGGCCTGGAATTGGCCGAC 559
QY      158 GlyGlnValTyr-----TrpValThrMetGlyThrAsnSerAlaLeuLysSerVal 174
Db      558 GGCAGCGCCCTGAGCTATGAATACGACGCCCTTGGCGGCAGACCCGTATC-----GCC 505
QY      175 ThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAla 194
Db      504 GATGCCGAGGGCACGCCACGCTGTTTTCCTGGGGGCACGGCGAT-----CTGCTGGCG 451
QY      195 ThrLysSerAsnGluAsnGlyTrpThrPheTyrGluTyrAspSerPheGlyArgLeu 214
Db      450 CGGGTTTCGACGCCGCGCGCGCGGAATTGCTCTACCTGCACGACGAAGCCGCGAGGCTG 391
QY      215 ThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAsp 230
Db      390 GTCGCCCTGACCAACGAGAACGGCGTCCAGGCGCAGTTCGCTACGAC 343

RESULT 14
US-09-252-991A-14461
; Sequence 14461, Application US/09252991A
```

```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14461
; LENGTH: 3387
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14461

Alignment Scores:
Pred. No.:          0.0864          Length:          3387
Score:             97.50          Matches:         30
Percent Similarity: 42.24%        Conservative:   19
Best Local Similarity: 25.86%     Mismatches:    60
Query Match:       7.04%          Indels:         7
DB:                  4            Gaps:           3

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-252-991A-14461 (1-3387)
QY      118 LeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnGly 137
Db      1513 GTCTGGCGTTACGACTACGACAAACGAGGACGTTTGGTCGCCAGCAGCCCGCTGGGA 1572
QY      138 AsnMetValAsnValArgAspSerThrGlyMetProLeuTrpLeuValValProAsp 157
Db      1573 CAACTGACCCGCGCGCTACGACCCGCTGGGCAACTGATCGGCCTGGAATTGGCCGAC 1632
QY      158 GlyGlnValTyr-----TrpValThrMetGlyThrAsnSerAlaLeuLysSerVal 174
Db      1633 GGCAGCGCCCTGAGCTATGAATACGACGCCCTTGGCGGCAGACCCGTATC-----GCC 1686
QY      175 ThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAla 194
Db      1687 GATCCGAGGGGCACGCCACGCTGTTTTCCTGGGGGCACGGCGAT-----CTGCTGGCG 1740
QY      195 ThrLysSerAsnGluAsnGlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeu 214
Db      1741 CGGGTTTCGACGCCGCGCGCGGCGGAATTGCTCTACCTGCACGACGAAGCCGCGAGGCTG 1800
QY      215 ThrAsnValThrPheProThrGlyGlnValSerSerPheArgSerAsp 230
Db      1801 GTCGCCCTGACCAACGAGAACGGCGTCCAGGCGCAGTTCGCTACGAC 1848

RESULT 15
US-09-221-017B-482/c
; Sequence 482, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 482:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...2236
US-09-221-017B-482

Alignment Scores:
Pred. No.: 0.0611 Length: 2236
Score: 96.50 Matches: 65
Percent Similarity: 34.12% Conservative: 22
Best Local Similarity: 25.49% Mismatches: 85
Query Match: 6.97% Indels: 83
DB: 4 Gaps: 14

US-10-029-020-14_COPY_1490_1750 (1-261) x US-09-221-017B-482 (1-2236)

QY 47 LeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGluLeuTyrValAlaAspLeu 66
Db 1907 TTGGAAGCTCCCCACCCTTTTCGGTG---GTCGAAGTTCATCGTAGAATGGCTGATCTC 1851
QY 67 GlyAsnIleArgIleArgPhe-----IleArg 75
Db 1850 GGCCGAATTGCGAAGCGGTTCTTTTCGTAATTACGTTGATAACGCCTCCACAGCATTAGA 1791
QY 76 LysAsnLysProPheLeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGln 95
Db 1790 GCCGAACAGACCGGAACCTCCACCGCAATTACTTC----- 1755
QY 96 GluLeuTyrLeuPheAspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGly 115
Db 1754 -----TACACGTCGATCATATTGGCAGGCATCTG---CTCCAAGCCATAGACACCGGC 1704
QY 116 AspTyrLeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsn 135
Db 1703 -----AAGCGAACTGAAGATGGG-----ATGGCTGTCGATAAG 1671
QY 136 Asn-----GlyAsnMetValAsnValArgAspSerThrGlyMetProLeuTyr 152
Db 1670 AATTTCGGAATAGGCTCCTTCGAGTCCATTGATACGAACTTGGTTGAAACCGCAGTTCTG 1611

QY 153 LeuValValProAspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAlaLeuLys 172
Db 1610 ACAGTTGTCTCCACGCGCAG-----ACCGGGCTGGAACCTAAG 1572
QY 173 SerValThrThrGlnGly-----HisGluLeuAlaMetMetThrTyrHisGly 188
Db 1571 TCCCTGACTCAGGTTGGTAGAGTTGGTTTGTAGGAAAAGTTCGGCGCAGACATACCGTTAC 1512
QY 189 AsnSerGlyLeuLeuAlaThrLysSer----- 197
Db 1511 CAACGAGGGTGCTTGACGGCGGAAAAGTCTCGTTTCTATTGGCAGATACCACTACCCCATC 1452
QY 198 ---AsnGluAsn-----GlyTrpThrThrPheTyr 206
Db 1451 GAGAGAGAGAATCTCTTCGGTCAGATAGAGGTCTACGTGGCGGGAACCACTCTTCTAT 1392
QY 207 GluTyrAspSerPheGlyArgLeuThrAsnValThrPheProThrGlyGlnValSer--- 225
Db 1391 ATGTACTTTTCTTCGCA-----AGTGGAGTAACCCATGAGTCGGGCTACGATAGTTTG 1338
QY 226 ---SerPheArgSerAspThrAspSerSerValHisValGlnValGluThrSerSerLys 244
Db 1337 ATTACCCGAAGGGATTCCCTTGATCGAGTAGTTTCCGTTTGATCCGTTGCGGAGTCCGAG 1278
QY 245 AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 1277 -----AGTGGTTCTCTTT 1266

Search completed: August 14, 2004, 21:37:30
Job time : 323.216 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 3060.96 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750
Perfect score: 1384
Sequence: 1 KINRIRQVTTSGEISLVAGA.....SSKDDVTITNLSASGAFYT 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_13135@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1384	100.0	6246	29	AY413475	AY413475 Homo sapi
2	1379	99.6	5970	29	AY413476	AY413476 Pan trogl
3	1370	99.0	3190	29	AY413477	AY413477 Mus muscu
4	951	68.7	728	14	CB520016	CB520016 UI-M-GIO-
5	943.5	68.2	5087	29	AY405422	AY405422 Mus muscu
6	925.5	66.9	5094	29	AY405420	AY405420 Homo sapi
7	865	62.5	759	13	BU704105	BU704105 UI-M-FOO-
8	857	61.9	723	13	BU704133	BU704133 UI-M-FOO-
9	847.5	61.2	737	14	CB520473	CB520473 UI-M-GIO-
10	830.5	60.0	797	14	CD349851	CD349851 UI-M-FY0-
11	812.5	58.7	5069	29	AY405421	AY405421 Pan trogl
12	791	57.2	739	13	BQ769387	BQ769387 UI-M-FIO-
13	790	57.1	735	29	AG045907	AG045907 Pan trogl
14	720.5	52.1	764	14	CB244782	CB244782 UI-M-FY0-
15	704	50.9	588	28	AZ625513	AZ625513 1M0465104
16	684.5	49.5	648	13	BU057564	BU057564 UI-M-FOO-
17	653.5	47.2	1494	28	CC206192	CC206192 CH261-170
18	649	46.9	602	9	AL654107	AL654107 AL654107
19	637	46.0	575	12	BG800702	BG800702 0052-06 M
20	593	42.8	786	14	CA752501	CA752501 UI-M-FOO-
21	575.5	41.6	551	12	BI975935	BI975935 484513 MA
22	570	41.2	793	13	BU205083	BU205083 603104346
23	568	41.0	512	29	CE005729	CE005729 tigr-gss-
24	549	39.7	472	14	CF531401	CF531401 UI-M-FY0-
25	546	39.5	782	14	CF537442	CF537442 UI-M-GIO-
26	541.5	39.1	789	14	CD349252	CD349252 UI-M-FY0-
27	532.5	38.5	942	29	CNS022KB	AL221060 Tetraodon
28	521	37.6	774	29	AG082821	AG082821 Pan trogl
29	497	35.9	716	14	CD803068	CD803068 UI-M-GV0-
30	496	35.8	794	14	CB518529	CB518529 UI-M-GH0-
31	492	35.5	854	14	CF745232	CF745232 UI-M-GV0-
32	443.5	32.0	730	9	AU051196	AU051196 AU051196
33	355	25.7	933	29	CNS041MM	AL270391 Tetraodon
34	342	24.7	399	28	AZ831924	AZ831924 2M0111N20
35	339.5	24.5	909	10	BF980526	BF980526 602304274
36	333	24.1	798	10	BF144774	BF144774 601791454
37	317.5	22.9	806	14	CF539403	CF539403 UI-M-GH0-
38	317	22.9	524	28	AZ005582	AZ005582 RPCI-23-3
39	317	22.9	601	13	BX509644	BX509644 DKFZp686N
40	316	22.8	1083	28	CC239138	CC239138 CH261-171
41	297	21.5	463	9	AL928416	AL928416 AL928416
42	295	21.3	545	10	BF075317	BF075317 224264 MA
43	295	21.3	900	13	BU267121	BU267121 603506284
44	289	20.9	1071	29	CNS059C0	AL327033 Tetraodon
45	286	20.7	602	9	AL710647	AL710647 DKFZp686M

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AY413475
Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY413475
AY413475.1 GI:39769437
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6246)

AY413475
Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY413475
AY413475.1 GI:39769437
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6246)

AY413475
Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY413475
AY413475.1 GI:39769437
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6246)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 6246)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..6246 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..>6246 /locus_tag="HCM4903"
gene
ORIGIN
Alignment Scores:
Pred. No.: 2.6e-154 Length: 6246
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AY413475 (1-6246)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 2398 AAGATCAACCGCATCAGGCGAGTCAACACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 2457
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 2458 CCCAGTGGCTGTGACTGTAAAAATGATGCGCAACTGTGATTGTTTCTTGAGACGATGGT 2517
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 2518 TATGCCAAGGATGCAAGTTAATAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGGAG 2577
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 2578 CTCACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGAAGAACAAGCCTTTC 2637
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGlnLeuTyrLeuPhe 100
Db 2638 CTCACACACCAGAACATGTATGAGCTGTCTTCACCAATTGACCAGGAGCTCTATCTGTTT 2697
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 2698 GATACCAACCGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 2757
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 2758 TTCACCTACACTGGGGACGGCGACATCACTACATCATCAGACAAATGGCAACATGGTA 2817
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 2818 AATGTCCCGCGAGACTCTACTGGATGCCCCCTCTGGCTGGTGGTCCAGATGGCCAGGTG 2877
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 2878 TACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCACACAAGGACACGAG 2937

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 2938 TTGGCCATGATGACATACCATGGCAATTCGGCCTTCTGGCAACCAAAAGCAATGAAAC 2997
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 2998 GGATGGACAAACATTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTGACCTTCCCT 3057
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 3058 ACTGGCCAGGTGACAGTTTCCGAAGTGATACAGACAGTTTCAGTGATGTCAGGTAGAG 3117
QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 3118 ACCTCCAGCAAGGATGATGTACCATTAACCAACCAACCTGTCTGCCTCAGGCGCTTCTAC 3177
QY 261 Thr 261
Db 3178 ACA 3180
RESULT 2
AY413476
LOCUS
DEFINITION Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY413476
VERSION AY413476.1 GI:39769438
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 5970)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5970)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..5970 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>5970 /locus_tag="HCM4903"
gene
ORIGIN
Alignment Scores:
Pred. No.: 9.67e-154 Length: 5970
Score: 1379.00 Matches: 260
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.62% Mismatches: 0
Query Match: 99.64% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AY413476 (1-5970)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 2122 AAGATCAACCGTATCAGGCGAGTCAACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGCC 2181

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 2182 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATGTTTTCTGGAGACGATGGT 2241
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 2242 TATGCCAAGGATGCAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGTGATGGGGAG 2301
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 2302 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTC 2361
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 2362 CTCAACACCCAGAACATGTATGAGCTGTCTTCACCAATTGACGAGGCTCTATCTGTTT 2421
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 2422 GATACCACTGGCAAGCACCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAAC 2481
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 2482 TTCACCTACACTGGGGACCGGTGACATCACACTCATCAGACAAACAAATGGCAACATGGTA 2541
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGlnVal 160
Db 2542 AATGTCGCCCGAGACTTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTG 2601
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 2602 TACTGGGTGACCATGGGCACCAACAGTGCACCTCAAGAGTGTGACCACACAAAGACACGAG 2661
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 2662 TTGGCCATGATGACATACCATGGCAATTCTGGCCTTCTGGCAACCAAAAGCAGTGAAAC 2721
QY 201 GlyTipThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 2722 GGATGGACAACATTTTATGAGTACGACAGCITTGGCCGCTGACAAATGTGACCTTCCCT 2781
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 2782 ACTGGCCAGGTGACGAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCAGGTAGAG 2841
QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 2842 ACCTCCAGCAAGGATGATGTACCATAACCAACCACTGTCTGCCTCAGGTGCCTTCTAC 2901
QY 261 Thr 261
Db 2902 ACA 2904
RESULT 3
AY413477
LOCUS 3190 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY413477
VERSION AY413477.1 GI:39769439
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3190)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)

14671302
2 (bases 1 to 3190)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
DIRECT SUBMISSION
SUBMITTED (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
Location/Qualifiers
Source
1..3190
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
gene
<1..>3190
/locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.: 4.2e-153 Length: 3190
Score: 1370.00 Matches: 258
Percent Similarity: 99.62% Conservative: 2
Best Local Similarity: 98.85% Mismatches: 1
Query Match: 98.99% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AY413477 (1-3190)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluLeuSerLeuValAlaGlyAla 20
Db 2398 AAGATCAATCGCATCAGCAGGTCACTACAGTGGTGAGATCTCACTGGTTGCTGGTGCC 2457
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 2458 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGACTGCTTCTCTGGAGATGATGGT 2517
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 2518 TACGCCAAGGATGCAAAAGCTGAATACCCCATCGTCTTGGCTGTGTGCTGACGGGGAG 2577
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 2578 CTCTATGTGGCCGACCTGGGAAACATCCGAATTCGATTTATCCGGAAGAACAAAGCCTTTC 2637
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 2638 CTGAACACTCAGAACATGTACGAGCTATCTCCCCATCGACCCAGGAGCTGTACCTCTT 2697
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 2698 GATACCACTGGCAAGCATCTGTACACTCAGAGCCTACCCACAGGGACTACCTGTACAAC 2757
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 2758 TTCACCTTACACAGGGGACGGGGACATCACATATCACCCGACAAATGGCAACATGGTG 2817
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGlnVal 160
Db 2818 AACGTCCGCCGAGACTCTACCCGGATGCCCTCTCTGGCTGGTAGTCCAGATGGCCAGGTA 2877
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 2878 TACTGGTTAACCATGGGCACCAACAGCGCACTCAGAAAGTGTGACCACACAAAGGACACGAG 2937
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 2938 CTAGCCATGATGACCTACCATGGCAACTCTGGCCTCTTGGCAACCAAAAGCAATGAAAAAC 2997
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 2998 GGTGGACAACGTTTTTATGATGATGACAGTTTTTGGTCGCCTGACAAACGTTGACCTTTTCCA 3057

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
|||
db 3058 ACTGCCCAGGTGAGCAGTTTCCGAAGCGATACAGACAGCTCAGTGCACGTGCAGGTAGAG 3117

RESULT 4	CB520016	LOCUS	CB520016	728 bp	mRNA	linear	EST 09-JUL-2003
DEFINITION	UI-M-GIO-cek-i-10-0-UI.r1 NIH_BMAP_GIO Mus musculus cDNA clone						
IMAGE:6840755 5', mRNA sequence.							

ORIGIN

Alignment Scores:	
Pred. No.:	1.13e-103
Score:	951.00
Percent Similarity:	99.45%
Best Local Similarity:	98.35%
Mismatches:	1
Conservative:	2
Matches:	179
Length:	728

Query Match:	68.71%	Indels:	0
DB:	14	Gaps:	0
US-10-029-020-14_COPY_1490_1750 (1-261) x CB520016 (1-728)			
QY	80	PheLeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeu	99
Db	7	TTCCTGAACACTCAGAACATGTACGAGCTATCTCTCCCCCATCGACCAGGAGCTGTACCTC	66
QY	100	PheAspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyr	119
Db	67	TTTGATACCAGTGGCAAGCATCTGTACACTCAGAGCCTACCCACAGGGACTACCTGTAC	126
QY	120	AsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMet	139
Db	127	AACTTCATTACAGGGGACGGGACATCACACATATCACCGACAACAATGGCAACATG	186
QY	140	ValAsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGln	159
Db	187	GTGAACGTCGCGCGAGACTCTACCGGATGCCCTCTCTGGCTGGTAGTCCCAGATGGCCAG	246
QY	160	ValTyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHis	179
Db	247	GTATACTGGGTAACCATGGGCACCAACAGCGCACTCAGAAAGTGTGACCACACAAGGACAC	306
QY	180	GluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGlu	199
Db	307	GAGCTAGCCATGATGACCTACCATGGCAACTCTGGCCCTCTGGCAACCAAAAGCAATGAA	366
QY	200	AsnGlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhe	219
Db	367	AACGGGTGGACCAACGTTTTATGAGTATGACAGTTTGGTCCGCTGACAAACGTGACCTTT	426
QY	220	ProThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnVal	239
Db	427	CCAACTGGCCAGGTGAGCAGTTTCCGAAGCGATACAGACAGCTCAGTGCACGTGCAGGTA	486
QY	240	GluThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe	259
Db	487	GAGACCTCAAGCAAGATGACGTACCATAACCAACCAACCTGTCTGTCTCGGGTGCCTTC	546
QY	260	TyrThr	261
Db	547	TACACC	552

RESULT 5	AY405422	LOCUS	AY405422	5087 bp	DNA	linear	GSS 12-DEC-2003
DEFINITION	Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.						
ACCESSION	AY405422	VERSION	AY405422.1	GI:39761396	GSS.		
KEYWORDS	Mus musculus (house mouse)						
SOURCE	Mus musculus						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 5087)						
REFERENCE	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.						
AUTHORS	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios						
TITLE	Science 302 (5652), 1960-1963 (2003)						
JOURNAL	14671302						
PUBMED	2 (bases 1 to 5087)						
REFERENCE	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.						
AUTHORS	Direct Submission						
TITLE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,						
JOURNAL							

Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
 source
 1..5087
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>5087
 /locus_tag="HCM2218"
ORIGIN
Alignment Scores:
Pred. No.: 2.2e-101 Length: 5087
Score: 943.50 Matches: 173
Percent Similarity: 79.39% Conservative: 35
Best Local Similarity: 66.03% Mismatch: 53
Query Match: 68.17% Indels: 1
DB: 29 Gaps: 1
US-10-029-020-14_COPY_1490_1750 (1-261) x AY405422 (1-5087)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 1561 AAGATCAACCGAATAAGCGAGTCCAGACAGACGGGGAGATCTCCTTAGTGGCTGGGATA 1620
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 1621 CCTTCGGAATGTGACTGCAGAAACGACGCGCAACTGTGACTGTACTACCAAGCGGAGACGGC 1680
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 1681 TAGCCCAAAGATGCCAAACTCAATGCGCGCTCCTCCCTGGCCGCTCGCCAGATGGCACT 1740
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 1741 CTGTACATTCAGATCTGGGAAATATCAGGATCCGGGCGGTTTCGAAGAATAAACCTTTA 1800
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 1801 CTGAACCTCAATGAACCTTTTACGAAGTTGCCTCTCCAACTGATCAAGAGCTCTACATCTTT 1860
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 1861 GACATCAACGGTACTCACCAGTACACCGTGAGCTGGTGCACGGGTGACTACTATATAAT 1920
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 1921 TTTAGTTACAGCAATGACAATGACGTACCGCTGTAACTGACAGCAATGGCAACACCCCTC 1980
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 1981 CGAATCCGAAGGATCCGAATCGGATGCCGGTGGGTGGTGTCTCCTGATAACCAAGGTG 2040
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 2041 ATATGGTTGACCATAGGCACCAACGGGTGTCTGAAAAGCATGACCGCTCAGGGCCTGGAA 2100
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 2101 CTGGTTTTTGTACTTACCATGGCAACAGTGGCTTTTAGCCACCAAAAGTGACGAAACT 2160
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 2161 GGATGGACAACATTTTGTACTATGACAGTGAAGGTGCGCTGACCAATGTTACCTTCCC 2220
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 2221 ACTGGGTGTTTACAAACCTGCACGGGGACATGGACAAGGCTATACCGTGGACATCGAG 2280
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 2281 TCATCCAGCAGAGAGGAAGATGTTCAGCATCACTTCGAACCTTGCTCCTCCATCGATTCTTC 2340

QY 260 TyrThr 261
Db 2341 TACACC 2346
RESULT 6
AY405420 5094 bp DNA linear GSS 12-DEC-2003
LOCUS Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
ACCESSION AY405420
VERSION AY405420.1 GI:39761394
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5094)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5094)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
FEATURES
 source
 1..5094
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>5094
 /locus_tag="HCM2218"
 gene
ORIGIN
Alignment Scores:
Pred. No.: 3.25e-99 Length: 5094
Score: 925.50 Matches: 172
Percent Similarity: 78.24% Conservative: 33
Best Local Similarity: 65.65% Mismatch: 56
Query Match: 66.87% Indels: 1
DB: 29 Gaps: 1
US-10-029-020-14_COPY_1490_1750 (1-261) x AY405420 (1-5094)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 1568 AAAATTAAACCGGATAAGCGAGGTCAACACAGATGGAGAAATCTCTTAGTGGCCGAATA 1627
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 1628 CCTTCAGAGTGTGACTGCAAAAATGATGCCAACTGTGACTGTGTACCAAGATGGAGATGGC 1687
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 1688 TAGCCCAAGGATGCCAACTCAGTGCCTCCATCTCCTGGTGTCTCTCCAGATGGTACA 1747
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 1748 CTGTATATTGCAGATCTAGGGAATATCCGATCCGGCTGTGTCAAAAGATAAGCCTTTA 1807
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

```
Db      1808 CTTAACTCTATGAACCTTCTATGAAGTTGGCGTCTCCAAGTATGATCAAGAACTCTACATCTTT 1867
QY      101  AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db      1868 GACATCAATGGTACTCACCATATACTGTAAGTTTAGTCACTGGTGATACCTTTTACAAT 1927
QY      121  PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db      1928 TTTAGCTACGCAATGACAATGATATATTACTGCTGTGACAGACAGCAATGGCAACACCCTT 1987
QY      141  AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db      1988 AGAATTAGCGGACCCAAATCGCATGCCAGTTCGAGTGGTGTCTCTCTGATAACCAAGTG 2047
QY      161  TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db      2048 ATATGGTTGACAATAGGAACAATGGATGTTTGAAGAGCATGACTGCTCAAGGACTGGAA 2107
QY      181  LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db      2108 TTAGTTTGTGTTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAGTGATGAAACT 2167
QY      201  GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db      2168 GGATGGACAACGTTNNNNNNCTATGACAGTGAAGTTCGTCTGACAAATGTTAGCTTTCCA 2227
QY      221  ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db      2228 ACTGGAGTGGTCACAAACCTGCATGGGACATGGACAGGCTATCACAGTGGACATTGAG 2287
QY      241  ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db      2288 TCATCTAGCCGAGAGAAGATGTCAGCATCACTTCAAATCTGTCTCGATCGATTCTTTC 2347
QY      260  TyrThr 261
Db      2348 TACACC 2353

RESULT 7
BU704105
LOCUS
DEFINITION
  BU704105 759 bp mRNA linear EST 15-JUL-2003
  UI-M-FOO-bzr-d-20-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
  IMAGE:6406243 5', mRNA sequence.
ACCESSION
  BU704105
VERSION
  BU704105.1 GI:23631890
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 759)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-re@mail.nih.gov
  Tissue Procurement: Dr. Jim Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
  1..759
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6"
  /db_xref="taxon:10090"
  /clone="IMAGE:6406243"
```

FEATURES
source

```
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_F00"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I;
Site_2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TGAGAGAGCC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2.77e-93 Length: 759
Score: 865.00 Matches: 161
Percent Similarity: 78.51% Conservative: 29
Best Local Similarity: 66.53% Mismatches: 52
Query Match: 62.50% Indels: 1
DB: 13 Gaps: 0
```

US-10-029-020-14_COPY_1490_1750 (1-261) x BU704105 (1-759)

```
QY      1  LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db      35  AAGATCAACCGAATAAGGAGGTCCAGACAGACGGGAGAGATCTCCTTAGTGGCTGGGATA 94
QY      21  ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db      95  CCTTCGGAATGTGACTGCAAGAACGACGCCAACTGTGACTGTACCAAAGCGGAGACGSC 154
QY      41  TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db      155  TAGGCCAAAGATGCCAAACTCAATGCGCGTCTCTCCCTGGCCGCTCGCCAGATGGCACT 214
QY      61  LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db      215  CTGTACATTGCAGATCTGGGAAATATCAGGATCCGGGCGCTTTCGAAGAATAAACCTTTA 274
QY      81  LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db      275  CTGAACCTCAATGAACCTTTACGAAGTTGCCTCTCCAACTGATCAAGAGCTCTACATCTTT 334
QY      101  AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db      335  GACATCAACGGTACTCACCCAGTACACCGTGAGCCTGGTCACGGGTGACTACCTATATAAT 394
QY      121  PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db      395  TTTAGTTACAGCAATGACAATGACGTCACCGCTGTAACCTGACAGCAATGGCAACACCCCTC 454
QY      141  AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db      455  CGAATCCGAAGGGATCCGAATCGATCCGATCCGGTGGTGTCTCTGATAACCCAGGTG 514
QY      161  TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db      515  ATATGGTTGACCATAGGCACCACCGGTGTCTGAAAAGCATGACCGCTCAGGGCCTGGAA 574
QY      181  LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db      575  CTGTTTGTGTTTACTTACCATGGCAACAGTGGGCTTTTAGCCACCANAAGTGACGAAACT 634
QY      201  GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
```


Db	635	GGATGGACAACATTTTGGACTATGACAGTGAAGGTGGCTGACGAATGTTACCTTCCCC	694
QY	221	ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu	240
Db	695	ACTGGGTGTTACAAACTG-CACGGGGACATGGACAAGGCTATCACGGTGGACATCGAG	753
QY	241	ThrSer	242
Db	754	TCATCC	759
RESULT 8			
BU704133			
LOCUS	BU704133	723 bp	mRNA
DEFINITION	UI-M-FO0-bzr-j-10-0-UI.r1 NIH_BMAP_F00	Mus musculus	CDNA clone
ACCESSION	IMAGE:6406377	5', mRNA sequence.	
VERSION	BU704133		
KEYWORDS	EST.		
SOURCE	BU704133.1	GI:23631946	
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
JOURNAL	1 (bases 1 to 723)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov This clone was contributed by the Brain Molecular Anatomy Project (BMAP)		
Seq primer:	pyX-5.		
FEATURES	Location/Qualifiers		
source	1. .723		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:6406377"		
	/tissue_type="whole brain"		
	/dev_stage="embryo 12.5dpc"		
	/lab_host="DH10B (T1 phage resistant)"		
	/clone_lib="NIH_BMAP_F00"		
	/note="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAGAGAGCC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."		
ORIGIN			
Alignment Scores:			
Pred. No.:	2.36e-92	Length:	723
Score:	857.00	Matches:	156
Percent Similarity:	80.09%	Conservative:	25
Best Local Similarity:	69.03%	Mismatches:	45
Query Match:	61.92%	Indels:	0
DB:	13	Gaps:	0

US-10-029-020-14_COPY_1490_1750 (1-261) x BU704133 (1-723)	
QY	1
Db	34
QY	21
Db	94
QY	41
Db	154
QY	61
Db	214
QY	81
Db	274
QY	101
Db	334
QY	121
Db	394
QY	141
Db	454
QY	161
Db	514
QY	181
Db	574
QY	201
Db	634
QY	221
Db	694
RESULT 9	
CB520473	
LOCUS	UI-M-GIO-cej-m-11-0-UI.r1 NIH_BMAP_GIO
DEFINITION	Mus musculus
ACCESSION	CB520473
VERSION	CB520473.1
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	1 (bases 1 to 737)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

CB520473 737 bp mRNA linear EST 09-JUL-2003
UI-M-GIO-cej-m-11-0-UI.r1 NIH_BMAP_GIO Mus musculus CDNA clone
IMAGE:6840468 5', mRNA sequence.

Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers
1..737
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840468"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_G10"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 3.39e-91 Length: 737
Score: 847.50 Matches: 155
Percent Similarity: 78.19% Conservative: 35
Best Local Similarity: 63.79% Mismatches: 51
Query Match: 61.24% Indels: 2
DB: 14 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x CB520473 (1-737)

QY 21 ProSerGlyCysAspCysLysAsnAlaAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 6 CCTCGGAATGTGACTGCAAGAACGACGCGCACTGTGACTGTCTACCAAGCGGAGACGGC 65
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 66 TAGCCAAAGATGCCAAACTCAATGCGCGTCTCTCCGCGCTCGCCAGATGGCACT 125
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db 126 CTGTACATTGCAGATCTGGGAATATCAGGATCCGGCCGTTTCGAAGAATAAACCTTTA 185
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 186 CTGAACCTCAATGAACCTTTACGAAGTTGCCTCTCCCACTGATCAAGAGCTCTACATCTT 245
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 246 GACATCAACGGTACTACCAAGTACACCGTGAGCCTGGTCACGGGTGACTACCTATATAAT 305
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 306 TTAGTTACAGCAATGACAAATGACGTACCGCTGTAACTGACAGCAATGGCAACACCTC 365
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 366 CGAATCCGAAGGATCCGAATCGGATCGCGTGCGGGTGGTGTCTCCTGTATAACCAAGTG 425
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180

Db 426 ATATGTTGACCATAGGCACCAACGGGTGTCTGAAAAGCATGACCGCTCAGGGCCTGGAA 485
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 486 CTGGTTTGTCTTACTTACCATGGCAACAGTGGGCTTTTAGCCACCAAAAGTGACGAACT 545
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 546 GGATGGACAACATTTTGTACTATGACAGTGAAGTGCCTGACGAATGTTACCTTCCCC 605
QY 221 ThrGlyGln-ValSerSerPheArgSerAspThrAspSerSerValHisValGlnValG1 240
Db 606 ACTGGNGTGGTTACAAACCTGCACGGGACATGGACAAGGCTATCACGGTGGACATCGA 665
QY 240 uThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPh 259
Db 666 GTCATCCAGCAGAGAGGAAGATGTCAGCATCACTTCGAACTTGTCCTCCATCGATTCTT 725
QY 259 eTyrThr 261
Db 726 CTACACC 732

RESULT 10
CD349851
LOCUS
DEFINITION
UI-M-FY0-cfr-n-12-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6855085 5', mRNA sequence.
CD349851
CD349851.1 GI:31141366
EST.
KEYWORDS
SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 797)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1..797
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6855085"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National

FEATURES
source

Db 1988 AGAATTAGACGGGACCCAAATCGCATGCCAGTTCGAGTGGTGTCTCCGATAACCAAGTG 2047

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180

Db 2048 ATATGGTTGACAATAGGAACAAATGGATGTTTGAAAGCATGACTGCTCAAGGACTGGAA 2107

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200

Db 2108 TTAGTTNNGTTTACTTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2167

QY 201 GlyTIpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220

Db 2168 NNN 2227

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240

Db 2228 ACTGGAGTGGTCACAAACCTGCATGGGACATGGACAAGGCTATCACAGTGGACATTGAG 2287

QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259

Db 2288 TCATCTAGCCGAGGAAGATGTCAGCATCACTTCAAATCTGCTCGATCGATTCTTTC 2347

QY 260 TyrThr 261

Db 2348 TACACC 2353

RESULT 12

BQ769387

LOCUS

DEFINITION

UI-M-F10-byq-i-13-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone

IMAGE:6400980 5', mRNA sequence.

BQ769387

BQ769387.1 GI:21977861

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 739)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pyX-5.

Location/Qualifiers

1. 739

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:6400980"

/tissue_type="whole brain"

/dev_stage="embryo 12.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_lib="NIH_BMAP_F10"

/note="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned

FEATURES

source

directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 2.19e-84 Length: 739

Score: 791.00 Matches: 149

Percent Similarity: 78.44% Conservative: 22

Best Local Similarity: 68.35% Mismatches: 47

Query Match: 57.15% Indels: 1

DB: 13 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x BQ769387 (1-739)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20

Db 86 AAGATCAACCGAATAAGGCAGGTACGACACGACGGGAGATCTCCTTAGTGGCTGGGATA 145

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40

Db 146 CCTTCGGAATGIGACTGCAAGAACGACGCCNACTGTGACTGTCTACCAAGCGGAGACGGC 205

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60

Db 206 TACGCCAAAGATGCCAAACTCAATGCGCCGTCCTCCCTGGCCGCTCGCCAGATGGCACT 265

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80

Db 266 CTGTACATTGCAGATCTGGGAATATCAGGATCCGGCCCGTTTCGAAGAATAAACCTTTA 325

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

Db 326 CTGAACCTCAATGAACCTTTACGAAGTTGCCCTCTCCAACTGATCAAGAGCTCTACATCTT 385

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120

Db 386 GACATCAACGGTACTCACCAAGTACACCGTGAGCCCTGGTGCACGGGTGACTACCTATATNAT 445

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnGlyAsnMetVal 140

Db 446 TTTAGTTACAGCAATGACAAATGACGTCAACCGCTGTAACTGACAGCAATGGCAACACCTC 505

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160

Db 506 CGAATCCGAAGGGATCCGAATCGGATGCCGTGCGGGTGGTGTCTCTCTGATAACCAAGTG 565

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180

Db 566 ATATGGTTGACCATAGGCACCAACGGGTGTCTGAAAAGCATGACCGCTCAGGGCCCTGGAA 625

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200

Db 626 CTGGTTNTGTTTACTTTACCATGGCAACAGTGGGCTTTTAGCCACCAAAAGTGACGAAACT 685

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThr 218

Db 686 GGATGGAC-ACATTTTGTGACTATGACAGTGAAGGTGCGCCTGACGAATGTTACT 738

RESULT 13

AG045907

LOCUS

DEFINITION

AG045907

ACCESSION

AG045907.1 GI:16582799

KEYWORDS

SOURCE

ORGANISM

Pan troglodytes (chimpanzee)

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

735 bp DNA linear GSS 02-NOV-2001

Pan troglodytes DNA, clone: PTB-024P12.F, genomic survey sequence.

REFERENCE 1 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 735)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
source Location/Qualifiers
1..735
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-024P12.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Alignment Scores:
Pred. No.: 2.86e-84 Length: 735
Score: 790.00 Matches: 147
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.32% Mismatches: 0
Query Match: 57.08% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x AG045907 (1-735)
QY 60 GluLeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysPro 79
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
135 GAGCTCTACGTGGCGGACCTTGGGAACATCCGAATTCGGTTTATCCGGAACAAGCCT 194
QY 80 PheLeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeu 99
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
195 TTCTCTCAACACCAGAACATGTATGAGCTGTCTTCCACCAATTGACCAGGAGCTCTATCTG 254
QY 100 PheAspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyr 119
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
255 TTTGATACCACTGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTAC 314
QY 120 AsnPheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMet 139
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
315 AACTTCACCTACACTGGGACGGTGACATCACACTCATCACAGACAATGGCAACATG 374
QY 140 ValAsnValArgAspSerThrGlyMetProLeuTrpLeuValProAspGlyGln 159
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
375 GTAAATGTCGTCGAGACTCTACTGGGATGCCCTCTGGCTGTGTCGCCAGATGGCCAG 434
QY 160 ValTyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrGlnGlyHis 179
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
435 GTGTACTGGGTGACCATGGGCACCAACACTGCACCTCAAGAGTGTACCAACAAGGACAC 494
QY 180 GluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGlu 199
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||
495 GAGTTGGCCATGATGACATACCATGGCAATTCTGGCCTTCTGGCAACCAAAAGCAGTGA 554

QY 200 AsnGlyTrpThrThrPheTyrGlu 207
Db ||||||||||||||||||||||||||||||||||||||||||||||||
555 AACGGATGGACAACACATTTTATGAG 578
RESULT 14
CB2444782 764 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-FY0-cdr-1-01-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
DEFINITION IMAGE:6833522 5', mRNA sequence.
ACCESSION CB2444782
VERSION CB2444782.1 GI:28366426
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 764)
NIH-MGC http://mgc.nci.nih.gov/
REFERENCE National Institutes of Health, Mammalian Gene Collection (MGC)
AUTHORS Unpublished (1999)
TITLE Contact: Robert Strausberg, Ph.D.
JOURNAL Email: cgapbs-remail.nih.gov
COMMENT Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
source Location/Qualifiers
1..764
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6833522"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FY0"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 7.2e-76 Length: 764
Score: 720.50 Matches: 146
Percent Similarity: 72.33% Conservative: 37
Best Local Similarity: 57.71% Mismatches: 56
Query Match: 52.06% Indels: 16
DB: 14 Gaps: 2
US-10-029-020-14_COPY_1490_1750 (1-261) x CB2444782 (1-764)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db ||||||||||||||||||||||||||||||||||||||||||||||||
36 AAGATCAACCGCCTACGCCAAGTCACCACCAATGGAGAGATGCTCTTTAGCCGGGCG 95

```
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 96 GCCTCAGACTGTGACTGCAAAAACGATGTCAACTGCATCTGTACTCGGAGATGACGCT 155
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 156 TACGCCACGGACGCATCCTAAACTCGCCGCTCTCTTAGCCGTGGCTCCAGATGGCACC 215
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProphe 80
Db 216 ATCTACATTGCAGACCTTGGGAATATCCGGATCAGGGCGGTTCAGCAAAAATAAACCCGTT 275
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 276 CTTAACGCATTCAACACAGTATGAGGCTGCATCTCCGGGAGAACAGGAATTGTACGTGTT 335
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 336 AACGCTGATGTTATCCATCAGTACACTGTGAGTCTGGTGACTGGGGAGTACTGTACAAT 395
QY 121 PheThrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 396 TTCACATACAGCGCTGACAATGACGTACCGAGTTGATTGACAACAACGGGAATTCCTTA 455
QY 141 AsnValArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 456 AAGATCCGCGGGACAGCAGTGGCATGCCCGCCACCTGCTCATGCTGATAATCAGATT 515
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 516 ATCACCTTACCCTGGGCACCAATGGAGGCTCAAAGCGTTTCCACTCAGAACCTGGAG 575
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 576 CTGGCCTCATGACTTATGATGGGAACACTGGACTCTAGCCACCAAGAGTGATGAAACT 635
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 636 GGATGGAC-ACCTTTTATGACTATGACACGGAGGCGTCT---GACAATGTGACCCG-CCC 690
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerValHisValGlnValGlu 240
Db 691 ACGGGCGTAGTGACC-----ATCTGC 711
QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeu 253
Db 712 ACCGGGAATGGAGAATCTATCACCATGACATTGAGACTC 750
```

```
RESULT 15
AZ625513/c
LOCUS
DEFINITION
  AZ625513
  1M0465104F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
  clone UUGC1M0465104 F, genomic survey sequence.
ACCESSION
  AZ625513
VERSION
  AZ625513.1 GI:11747703
KEYWORDS
  GSS.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 588)
REFERENCE
  Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
  Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
  Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
  Niederhausern,A. and Wright,D.,Weiss,R.
  Mouse whole genome scaffolding with paired end reads from 10kb
  plasmid inserts
  Unpublished (2000)
JOURNAL
  Contact: Robert B. Weiss
  University of Utah Genome Center
  University of Utah
  Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
```

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0465 row: I column: 04
Seq primer: CGTTGTAACACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 588.

FEATURES
source

Location/Qualifiers
1..588
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0465104"
/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Alignment Scores:
Pred. No.: 4.56e-74 Length: 588
Score: 704.00 Matches: 132
Percent Similarity: 98.53% Conservative: 2
Best Local Similarity: 97.06% Mismatches: 1
Query Match: 50.87% Indels: 1
DB: 28 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x AZ625513 (1-588)

```
QY 73 PheIleArgLysAsnLysPropheLeuAsnThrGlnAsnMetTyrGluLeuSerSer-Pr 92
Db 586 TTATCCGGAAGAACAGCCTTCTCTGAACACTCAGAACATGTACGAGCTATCCTCCCCC 527
QY 92 oIleAspGlnGluLeuTyrLeuPheAspThrThrGlyLysHisLeuTyrThrGlnSerLe 112
Db 526 CATCGACCAGGAGCTGTACTCTTTGATACCAGTGGCAAGCATCTGTACACTCAGAGCCT 467
QY 112 uProThrGlyAspTyrLeuTyrAsnPheThrTyrThrGlyAspGlyAspIleThrLeuI 132
Db 466 ACCACAGGGGACTACCTGTACAACTTCACCTTACACAGGGGACGGGACATCACACATAT 407
QY 132 eThrAspAsnAsnGlyAsnMetValAsnValArgArgAspSerThrGlyMetProIleTr 152
Db 406 CACCGACAACAATGGCAACATGCTGTGAACGCTCCGCCGAGACTCTACCGGGATGCCTCTCTG 347
QY 152 pLeuValValProAspGlyGlnValTyrTrpValThrMetGlyThrAsnSerAlaLeuLy 172
Db 346 GCTGGTAGTCCAGATGGCCAGTATACTGGGTAAACCATGGGCACCAACAGCGCACTCAG 287
QY 172 sSerValThrThrGlnGlyHisGluLeuAlaMetMetThrTyrHisGlyAsnSerGlyLe 192
Db 286 AAGTGTGACCACACAAGGACACGAGCTAGCCATGATGACCTACCATGGCAACTCTGGCCT 227
```


QY 192 uLeuAlaThrLysSerAsnGluAsnGlyTrpThrThrPheTyrGlu 207
Db 226 CTTGGCAACCAAGCAATGAAAAACGGGTGGACAAACGTTTATGAG 181

Search completed: August 14, 2004, 18:03:46
Job time : 3083.96 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 442.726 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1490_1750
Perfect score: 1384
Sequence: 1 KINRIRQVTTSGEISLVAGA.....SSKDDVTITNLSASGAFYT 261

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1868@runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1384	100.0	8354	6 ABS52100	AbS52100 Human TEN
2	1384	100.0	8438	6 ABN85378	Abn85378 Human NOV
3	1380	99.7	8645	6 ABS78652	AbS78652 Human CDN
4	935.5	67.6	8473	6 ABQ82345	Abq82345 Human NOV
5	935.5	67.6	8487	6 ABQ82346	Abq82346 Human NOV
6	935.5	67.6	8645	6 ABQ82344	Abq82344 Human NOV
7	935.5	67.6	8675	6 ABQ82343	Abq82343 Human NOV
8	898	64.9	12879	6 ABK92230	Abk92230 Prostate

9	898	64.9	13202	4	AAK51828	Aak51828 Human pol
10	863.5	62.4	9058	7	ACC72051	Acc72051 BCU0205A
11	863.5	62.4	9695	7	ACC72052	Acc72052 BCU0205B
12	863.5	62.4	9729	5	AAS14089	Aas14089 Human FCT
13	863.5	62.4	9729	9	ADB32028	Adb32028 Human FCT
14	863.5	62.4	9826	5	AAS14085	Aas14085 Human FCT
15	863.5	62.4	9826	9	ADB32023	Adb32023 Human FCT
16	812.5	58.7	4245	7	AAL60066	Aal60066 Human PCO
17	522.5	37.8	2434	5	AAS71215	Aas71215 DNA encod
18	390.5	28.2	10242	4	ABL29075	AbL29075 Drosophil
19	390.5	28.2	17131	4	ABL29074	AbL29074 Drosophil
20	389	28.1	3310	6	ABK34713	Abk34713 Human CDN
21	270.5	19.5	776	4	AAH05747	Aah05747 Human CDN
22	270.5	19.5	3614	4	AAH14183	Aah14183 Human CDN
23	193	13.9	3910	4	ABL04849	AbL04849 Drosophil
24	193	13.9	6318	4	ABL04848	AbL04848 Drosophil
25	131.5	9.5	548	6	ABQ37165	Abq37165 Oligonucl
26	131.5	9.5	548	6	ABQ37164	Abq37164 Oligonucl
27	119.5	8.6	4716	7	ACA45361	Aca45361 Prokaryot
28	115	8.3	31960	8	ACD18998	Acd18998 E. coli 0
29	115	8.3	35910	9	ADC00933	Adc00933 Enterohae
30	112.5	8.1	2415	8	ADA30132	Ada30132 DNA encod
31	111.5	8.1	1235	6	ABQ68530	Abq68530 Listeria
32	110.5	8.0	5394	6	ABQ71024	Abq71024 Listeria
33	110	7.9	6567	4	AAS59636	Aas59636 Propionib
34	110	7.9	6567	7	ACF64565	Acf64565 Propionib
35	109.5	7.9	1764	4	ABL04855	AbL04855 Drosophil
36	109.5	7.9	4925	4	ABL04854	AbL04854 Drosophil
37	109.5	7.9	6392	6	ABQ67876	Abq67876 Listeria
38	109.5	7.9	6504	6	ABQ69406	Abq69406 Listeria
39	109.5	7.9	110000	6	ABQ69245_04	Continuation (5 of
40	109.5	7.9	110000	6	ABQ67197_03	Continuation (4 of
41	108	7.8	4155	7	ACA44188	Aca44188 Prokaryot
42	107.5	7.8	2532	7	ACA52160	Aca52160 Prokaryot
43	107.5	7.8	4065	7	ACA52159	Aca52159 Prokaryot
44	107	7.7	4497	7	ACF72220	Acf72220 Photorhab
45	107	7.7	96798	7	ACF65387_6	Continuation (7 of

ALIGNMENTS

RESULT 1
ABS52100

ID ABS52100 standard; DNA; 8354 BP.

XX ABS52100;

XX ABS52100;

DT 05-NOV-2002 (first entry)

XX Human TEN-M4-like gene.

Human; NOX; NOX-associated disorder; cardiomyopathy; atherosclerosis;
cell signal processing; metabolic pathway modulation; metabolic disorder;
obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
memory defect; infertility; congenital heart defect; hair growth;
pigmentation disorder; endocrine disorder; respiratory disease; health;
gastro-intestinal disease; reproductive; neurological disease;
bone marrow transplantation; endocrine disease; allergy; inflammation;
nephrological disorder; urinary system disorder; age-related disorder;
neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
adipocyte complement-related Clq tumour necrosis factor; out at first;
beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
type Ia membrane sushi-containing domain; butyrophilin;
single nucleotide polymorphism.

OS Homo sapiens.

XX Key Location/Qualifiers
FH variation replace(117,G)
FT /*tag= a

FT /standard name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
XX 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
PR 20-DEC-2000; 2000US-0257314P.
PR 02-MAY-2001; 2001US-0288153P.
PR 29-MAY-2001; 2001US-0294075P.
PR 24-JUL-2001; 2001US-0307506P.
PR 10-AUG-2001; 2001US-0311590P.
PR 10-AUG-2001; 2001US-0311613P.
PR 29-AUG-2001; 2001US-0315617P.
PR 14-SEP-2001; 2001US-0322358P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shinkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
FT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.
PT
XX Claim 8; Page 50-52; 318pp; English.
PS
XX The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders. The
CC present nucleic acid sequence represents a NOVX gene. This sequence
CC encodes a NOVX protein of the invention
XX
SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.84e-133 Length: 8354
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-029-020-14_COPY_1490_1750 (1-261) x ABS52100 (1-8354)
QY 1 LysIleAsnArgIleArgGlnValThrSerGlyGluIleSerLeuValAlaGlyAla 20
DB 4502 AAGATCAACCGCATCAGGCAGGTACCACCTAGTGGAGAGATCTCACTGTTGCTGGGGCC 4561

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4562 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTTCTGGAGACGATGGT 4621
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4622 TATGCCAAGGATGCAAAAGTTAAATACCCCACTCTTCTTGGCTGTGTGCTGATGGGAG 4681
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db 4682 CTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTATCCGGAAGAACAGCCTTTC 4741
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4742 CTCAACACCCAGAACATGTATGAGCTGTCTTACCAATTGACCAGGAGCTCTATCTGTTT 4801
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4802 GATACCACCGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 4861
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4862 TTCACCTACACTGGGACGGCGACATCACACTCATCAGACACAACAATGGCAACATGCTA 4921
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 4922 AATGTCCGCCGAGACTCTACTGGGATGCCCTCTGGTGGTGGTCCCAGATGGCCAGGTG 4981
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4982 TACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACCACACAAGGACACGAG 5041
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5042 TTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAC 5101
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5102 GGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCTGACAAATGTGACCTTCCCT 5161
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5162 ACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGCATGTCCAGGTAGAG 5221
QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5222 ACTCCAGCAAGGATGATGTACCATTAACCAACCTGTGTGCTCAGGGCCTTCTAC 5281
QY 261 Thr 261
Db 5282 ACA 5284
RESULT 2
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
XX
AC ABN85378;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human NOV1, TEN-M4 like protein, coding sequence.
XX
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 4. .8395
FT /*tag= a
FT /trans_except= (pos: 1138. .1147,aa:Met)
FT /product= "NOV1 protein"
XX
XX WO200255704-A2.
XX
XX 18-JUL-2002..
XX
XX 09-JAN-2002; 2002WO-US000554.
XX
XX 09-JAN-2001; 2001US-0260417P.
PR 10-JAN-2001; 2001US-0260831P.
PR 28-FEB-2001; 2001US-0272338P.
PR 09-MAR-2001; 2001US-0274876P.
PR 18-APR-2001; 2001US-0284704P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
PI Taylor S, Tchernev VI, Miller CE, Guo X, Boldog FL, Grosse WM;
PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
PI Gunther E, Stone DJ;
XX
XX WPI; 2002-590674/63.
DR P-PSDB; ABB98401.
XX
XX NOVX polypeptides and encoding polynucleotides, useful for preventing or
PT treating NOVX-associated disorders e.g. cancer, inflammation, or
PT Alzheimer's disease, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 9; Page 8-9; 358pp; English.
XX
XX The present sequence is a coding sequence for a NOV protein. The NOV
CC proteins and coding sequences are useful for treating or preventing NOV-
CC associated disorders or in the manufacture of a medicament for treating
CC the disorders, such as cancer, heart disease, inflammation, autoimmune
CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
CC IgA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
CC and other wasting disorders associated with chronic diseases. NOV1 is a
CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX
SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.91e-133 Length: 8438
Score: 1384.00 Matches: 261
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x ABN85378 (1-8438)

QY 1 LysileAsnArgilleArgGlnValThrThrSerGlyGlulieserLeuValAlaGlyAla 20
Db 4508 AAGATCAACCGCATCAGGCAGGTTCACCACTAGTGGAGAGATCTACTCGTTGGGGCC 4567

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db 4568 CCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTCTGGAGACGATGGT 4627

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4628 TATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGCTGATGGGGAG 4687

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4688 CTCTAGTGGCCGACCTTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTC 4747

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleLeuAspGlnGluLeuTyrLeuPhe 100
Db 4748 CTC AACACCCAGAACATGTATGAGCTGTCTTACCAATTGACCAAGAGCTCTATCTGTTT 4807

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4808 GATACACCGGCAAGCACCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAAC 4867

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4868 TTCACCTACACTGGGGACGGGACATCACACTCATCACAGACAACAATGGCAACATGGTA 4927

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4928 AATGTCCGCCGAGACTCTACTGGGATGCCCCCTCTGGCTGGTGGTCCCAGATGGCCAGGTG 4987

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4988 TACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACCACACAAGGACACGAG 5047

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5048 TTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAAC 5107

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5108 GGATGGACAACATTTATGAGTACGACAGCTTTGGCCGCTTGACAAATGTACCTTCCCT 5167

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5168 ACTGGCCAGGTGACAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCAGGTAGAG 5227

QY 241 ThrSerSerLysAspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5228 ACCTCCAGCAAGGATGATGTACCATAAACCACCAACCTGTCTGCCTCAGGCGCTTCTAC 5287

QY 261 Thr 261
Db 5288 ACA 5290

RESULT 3
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
AC ABS78652;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; infection;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200272830-A2.
XX
PD 19-SEP-2002.
XX

FT CDS 258..8144
FT /*tag= a
FT /product= "NOV15c"

PN WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

PR 29-DEC-2000; 2000US-0258928P.

PR 02-JAN-2001; 2001US-0259415P.

PR 04-JAN-2001; 2001US-0259785P.

PR 20-FEB-2001; 2001US-0269814P.

PR 09-MAR-2001; 2001US-0279863P.

PR 29-MAR-2001; 2001US-0279832P.

PR 13-APR-2001; 2001US-0279833P.

PR 18-APR-2001; 2001US-0284447P.

PR 25-APR-2001; 2001US-0286683P.

PR 29-MAY-2001; 2001US-0294080P.

PR 16-AUG-2001; 2001US-0312915P.

PR 17-AUG-2001; 2001US-0313325P.

PR 17-SEP-2001; 2001US-0322699P.

PR 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

PI Gunther E, Smithson G, Millet I, Macdougall JR;

DR WPI; 2002-732706/79.

DR P-PSDB; ABP53588.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.

PS Claim 8; Page 119-121; 444pp; English.

XX The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence encodes human NOV15c, which is located on chromosome 4

SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 2.34e-86 Length: 8473

Score: 935.50 Matches: 172

Percent Similarity:	79.01%	Conservative:	35
Best Local Similarity:	65.65%	Mismatches:	54
Query Match:	67.59%	Indels:	1
DB:	6	Gaps:	1
US-10-029-020-14_COPY_1490_1750 (1-261) x ABQ82345 (1-8473)			
QY	1	LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla	20
Db	4317	AAAATTAAACGGGATAAGGCAGGTCACAAACAGATGGAGAAATCTCTTAGTGGCCGGAATA	4376
QY	21	ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly	40
Db	4377	CCTTCAGAGTGTGACTGCAAAATGATGCCAACTGTGACTGTACAGATGGAGATGGC	4436
QY	41	TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu	60
Db	4437	TACGCCAAGGATGCCAACTCAGTCCCATCTCCCTGGCTGCTTCTCCAGATGGTACA	4496
QY	61	LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe	80
Db	4497	CTGTATATTGCAGATCTAGGGAATATCCGGATCCGGGCTGTGTCAAGAATAAGCCTTTA	4556
QY	81	LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe	100
Db	4557	CTTAACTCTATGAACCTTCTATGAAGTTGCGTCTCCACTGATCAAGAACTCTACATCTT	4616
QY	101	AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn	120
Db	4617	GACATCAATGGTACTACCAATATACTGTAAGTTAGTCACTGGTATTACCTTTACAAT	4676
QY	121	PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal	140
Db	4677	TTTAGCTACGCAATGACAATGAAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAA	4736
QY	141	AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal	160
Db	4737	AGAAATTAGACGGGACCAAAATGCGATGCCAGTTCCGAGTGGTGTCTCCTGATAACCAAGTG	4796
QY	161	TyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu	180
Db	4797	ATATGGTTGACAATAGGAACAAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAA	4856
QY	181	LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn	200
Db	4857	TTAGTTTGTGTTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAAGTGATGAAACT	4916
QY	201	GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro	220
Db	4917	GGATGGACACGTTTTTTGACTATGACATGACAGTGAAGGTCGTCTGACAAATGTTACGTTTCCA	4976
QY	221	ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu	240
Db	4977	ACTGGAGTGGTCACAAACCTGTCATGGGACATGGACAAGGCTATCACAGTGGACATTGAG	5036
QY	241	ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe	259
Db	5037	TCATCTAGCCGAGAAGAAGATGTGACGATCACTTCAAATCTGTCTCTCGATCGATTCTTTC	5096
QY	260	TyrThr 261	
Db	5097	TACACC 5102	
RESULT 5			
ABQ82346			
ID	ABQ82346	standard; cDNA; 8487 BP.	
XX	AC	ABQ82346;	
XX	DT	17-DEC-2002 (first entry)	
XX	DE	Human NOV15d encoding cDNA SEQ ID NO:41.	
XX			

Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; Pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; ss.

OS	Homo sapiens.
XX	
FX	Key
FT	CDS
FT	Location/Qualifiers
FT	299..8140
FT	/*tag= a
FT	/product= "NOV15d"

WO200262999-A2.

15-AUG-2002.

31-DEC-2001; 2001WO-US049976.

PR	29-DEC-2000;	2000US-0258928P.
PR	02-JAN-2001;	2001US-0259415P.
PR	04-JAN-2001;	2001US-0259785P.
PR	20-FEB-2001;	2001US-0269814P.
PR	09-MAR-2001;	2001US-0279863P.
PR	29-MAR-2001;	2001US-0279832P.
PR	29-MAR-2001;	2001US-0279833P.
PR	13-APR-2001;	2001US-0283889P.
PR	18-APR-2001;	2001US-0284447P.
PR	25-APR-2001;	2001US-0286683P.
PR	29-MAY-2001;	2001US-0294080P.
PR	16-AUG-2001;	2001US-0312915P.
PR	17-AUG-2001;	2001US-0313325P.
PR	17-SEP-2001;	2001US-0322699P.
PR	26-NOV-2001;	2001US-0333350P.

(CURA-) CURAGEN CORP.

xx
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangalli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther R, Smithson G, Millet I, Macdougall JR;

WPI; 2002-732706/79.
P-PSDB; ABP53589.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.

Claim 8: page 123-125; 444pp; English.

The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal

QY 260 TyrThr 261
|
Db 5093 TACACC 5098

RESULT 6
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.
XX
AC ABQ82344;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15b encoding cDNA SEQ ID NO:37.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; Pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric disorder; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; ss.

OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 151..8316
FT /*tag= a
FT /product= "NOV15b"
XX
PN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53587.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.
XX
PS Claim 8; Page 114-117; 444pp; English.

XX
CC The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence encodes human NOV15b, which is located on chromosome 4

XX
SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 2.4e-86 Length: 8645
Score: 935.50 Matches: 172
Percent Similarity: 79.01% Conservative: 35
Best Local Similarity: 65.65% Mismatches: 54
Query Match: 67.59% Indels: 1
DB: 6 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x ABQ82344 (1-8645)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
|
Db 4489 AAAATTAAACCGGATAGGAGGTGCACACAGATGGAGAAATCTCCTTAGTGGCGGAATA 4548

QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
|
Db 4549 CCTTCAGAGTGTGACTGCAAAAATGATGCCAAGTGTGACTGTTACCAGATGGAGATGGC 4608

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
|
Db 4609 TAGCCCAAGGATGCCAAACTCAGTGTCCCATCCTCCCTGGCTGCTTCTCCAGATGGTACA 4668

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
|
Db 4669 CTGTATATTGCAGATCTAGGGAATATCCGGATATCCGGATCCGGCTGTGTCAAAGAATAAGCCTTTA 4728

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
|
Db 4729 CTTAACTCTATGAACCTTCTATGAAGTTGCTCTCCAAGTATCAAGAACTCTACATCTTT 4788

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
|
Db 4789 GACATCAATGGTACTCACCATAATATCTGTAAAGTTTAGTCACGTGGTGATTACCTTTACAAT 4848

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
|
Db 4849 TTTAGCTACAGCAATGACAATGATATGATCTGTGTGACAGACGAATGGCAACACCTT 4908

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
|
Db 4909 AGAATTAGACGGGACCCAAATCGCATGCCAGTTCGAGTGGTGTCTCTGTATAACCAAGTG 4968

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
|
Db 4969 ATATGTTGACAATAGGAACAATGGATGTTTGAAGGCGTACTGCTCAAGGACTGGAA 5028

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4801 GACATCAATGGTACTACCAATATACTGTAAGTTTAGTCACCTGGTGATTACCTTTACAAT 4860
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4861 TTTAGCTACAGCAATGACAATGATATTACTGCTGTGACAGACAGCAATGGCAACACCCCTT 4920
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4921 AGAATTAGACGGGACCCAAATCGCATGCCAGTTCGAGTGGTGTCTCTGTATACCAAGTG 4980
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4981 ATATGGTTGACAATAGGAACAAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAA 5040
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5041 TTAGTTTGTGTTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAAGTGATGAAACT 5100
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5101 GGATGACAACCGTTTTTTTGACTATGACAGTGAAGTTCGTCTGACAAATGTTACGTTTCCA 5160
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5161 ACTGGAGTGGTCAACAACCTGCATGGGACATGGACAAGGCTATCACAGTGGACATTGAG 5220
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5221 TCATCTAGCCGAGAGAAGATGTTCAGCATCACTTCAAATCTGCTCGATCGATTCTTTC 5280
QY 260 TyrThr 261
Db 5281 TACACC 5286

RESULT 8

ABK92230
ID ABK92230 standard; DNA; 12879 BP.
XX
AC ABK92230;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated DNA sequence #116.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene; ds.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX WPI; 2002-471335/50.
DR

DR P-PSDB; ABG61913.

XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX

PS Claim 22; Page 394-397; 436pp; English.

XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
CC sequences

XX
SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3.16e-82 Length: 12879
Score: 898.00 Matches: 162
Percent Similarity: 78.46% Conservative: 42
Best Local Similarity: 62.31% Mismatches: 56
Query Match: 64.88% Indels: 0
DB: Gaps: 0

US-10-029-020-14_COPY_1490_1750 (1-261) x ABK92230 (1-12879)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4406 AAAGTAAACCGCATTTCAGCAAGTAACCAACCAATGGGAGATCTACATCATCGCTGGTGCC 4465
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4466 CCCACTGACTGTGACTGCAAAATTTGATCCAACTGTGACTGTTTTCAGGTGATGTGGC 4525
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4526 TATGCCAAAGATGCAAAAGATGAAAGCCCTTCCTCTTAGCAGTGTGCGCTGATGGAACC 4585
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProphe 80
Db 4586 CTCTATGTGGCAGACCTCGGAAATGTTTCGAATTCGTACCATCAGCAGGAACCAAGCCAC 4645
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuphe 100
Db 4646 CTGAATGACATGAACATTTATGAGATTGCTTCACCCGCTGATCAGGAACGTACCAAGTTC 4705
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4706 ACTGTAATGGAACCCACCTACACACCTGAACTTGATAACAAGGACTATGTTATAAC 4765
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4766 TTCACCTACAATTTCTGAAGGTGACTTGGCGCGGATTACCCAGCAGCAATGCAATTCAGTG 4825
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 4826 CACATTCGCCGTGATGCAGGCGGGAATGCCGCTATGGCTTGTGGTGCCTGGCGGACAAGTA 4885
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 4886 TACTGGCTGACTATAAGCAGCAATGGAGTCCCTGAAAAGAGTGTTCAGCCCAAGGCTATAAT 4945

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 4946 CCGGCCTTAATGACCTATCCAGGAACACAGGGCTTCTGGCTACCAAAAGTAACGAAAT 5005
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5006 GGATGGACAACCGTTATGAGTATGACCCCGAGGACACCTGACCAATGCAACGTTTCCC 5065
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5066 ACTGGAGAGGTCAGCAGCTTCCACAGTGACCTGACGAGAGCTGACAAAGTGGAGCTAGAT 5125
QY 241 ThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5126 ACTTCCAACCGTGAAATGTCCTCATGTCAACCACTTGACGGCAACTAGTACCATATAT 5185

RESULT 9
AAK51828
ID AAK51828 standard; cDNA; 13202 BP.
XX
AC AAK51828;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 373.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

DR WPI; 2001-476283/51.
DR P-PSDB; AAM78695.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PI in diagnosis and gene therapy.
XX
PS Claim 1; Page 1414-1426; 6221pp; English.

XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;

Alignment Scores: 3.27e-82 Length: 13202
Pred. No.: 898.00 Matches: 162
Score: 78.46% Conservative: 42
Percent Similarity: 62.31% Mismatches: 56
Best Local Similarity: 64.88% Indels: 0
Query Match: 4 Gaps: 0
DB:

US-10-029-020-14_COPY_1490_1750 (1-261) x AAK51828 (1-13202)
QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db 4732 AAAGTAAACCGCATTCAGCAAGTAACCAACCAATGGGAGATCTACATCATCGCTGGTGCC 4791
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4792 CCCACTGACTGTGACTGCAAAATTTGATCCAACTGACTGTTTTTTCAGGTGATGGTGGC 4851
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4852 TATGCCAAAGATGCAAAAGATGAAAGCCCTTCCTCTAGCAGTGTGCGCTGATGGAACC 4911
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4912 CTCTATGTGGCAGACCTCGAAATGTTTGAATTCGTACCATCAGCAGGAACCAAGCCAC 4971
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4972 CTGAATGACATGAACATTTATGAGATTGCTTCAACCCGCTGATCAGGAACGTGACCAGTTC 5031
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 5032 ACTGTAAATGGAAACCCACCTACACCCCTGAACCTTGATAACAAGGGACTATGTTTATAAC 5091
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 5092 TTCACCTACAATCTGAAGGTGACTTGGGGCGGATTTACCAGCAGCAATGGCAATTCAGTG 5151
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160
Db 5152 CACATTCGCCGTGATGTCAGGGGAATGCCCTATGGCTTGGTGGCTGGCGGACAAGTA 5211
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5212 TACTGGCTGACTATAAGCAGCAATGGAGTCTTGAAAGAGTGTGAGCCCAAGGTATAAT 5271
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5272 CCGGCCTTAATGACCTATCCAGGAACACACAGGGCTTCTGGCTACCAAAAGTAACGAAAT 5331
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5332 GGATGGACAACCGTTTATGAGTATGACCCCGAGGACACCTGACCAATGCAACGTTTCCC 5391
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5392 ACTGGAGAGGTCAGCAGCTTCCACAGTGACCTGGAGAGCTGACAAAGTGGAGCTAGAT 5451
QY 241 ThrSerSerLysAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPheTyr 260
Db 5452 ACTTCCAACCGTGAAATGTCCTCATGTCAACCACTTGACGGCAACTAGTACCATATAT 5511

RESULT 10
ACC72051
ID ACC72051 standard; DNA; 9058 BP.
XX
AC ACC72051;


```
XX 08-JUL-2003 (first entry)
XX
XX BCU0205A gene #SEQ ID 79.
DE Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
XX drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
KW Homo sapiens.
XX
XX WO2003029421-A2.
XX
XX 10-APR-2003.
XX
XX 02-OCT-2002; 2002WO-US031287.
XX
XX 03-OCT-2001; 2001US-0326526P.
XX 14-MAY-2002; 2002US-00144194.
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
PI WPI; 2003-381623/36.
XX P-PSDB; ABR58317.
XX
XX New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
XX Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
XX The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records,
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;
XX
XX Alignment Scores:
Pred. No.: 7.7e-79 Length: 9058
Score: 863.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.39% Indels: 1
DB: 7 Gaps: 1
XX
XX US-10-029-020-14_COPY_1490_1750 (1-261) x ACC72051 (1-9058)
XX
XX 1 LysilleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3863 AAGATTACCGTCTACGCCAGGTAAACAACCAACGGGGAGATCTGCCTTTTAGCTGGGCA 3922
XX
XX 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3923 GCCTCGGACTGCGACTGCAAAAACGATGTCATTGCAACTGCTATTTCAGGAGATGATGCC 3982
XX
XX 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3983 TACGCGACTGATGCCATCTTGAAATCCCATCATCCTTAGCTAGTCTCCAGATGGTACC 4042
XX
XX 61 LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProphe 80
XX ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX WPI; 2003-381623/36.
XX P-PSDB; ABR58318.
```

```
Db 4043 ATTTACATTGCAGACCTTGGAAATATTTCGGATCAGGGCGGTGAGCAAGCAAGCCTGTT 4102
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4103 CTTAATGCCTTCAACCAAGTATGAGGCTGCATCCCCGGAGAGCAGAGTTATATGTTTC 4162
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
XX ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4163 AACGCTGATGGCATCCACCAATACACTGTGAGCCTGGTGACAGGGGAGTACTTGTACAAT 4222
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4223 TTCACATATAGTACTGACAATGATGTCACTGAATTGACAAATAATGGGAATTCCCTG 4282
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
XX ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4283 AAGATCCGTCGGGACAGCAGTGGCATGCCCGTACCTGCTCATGCTGACCAACCATGATC 4342
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
XX ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4343 ATCACCCCTCACCGTGGGCAACCAATGGAGGCCTCAAAGTCTGTCTCCACACAGAACCTGGAG 4402
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
XX ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4403 CTGGTCTCATGACCTATGATGGCAACACTGGGCTCCTGGCCACCAAGAGCGATGAACA 4462
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4463 GGATGGACGACTTCTATGACTATGACCACGAGGCGCCTGACCAACGTGACGCGCCCC 4522
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
XX |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4523 ACGGGGGTGGTAACCAAGTCTGCACCGGGAATGGAGAAATCTATTACCATTGACATTGAG 4582
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
XX ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4583 AACTCCAACCGTGATGATGACGTCACTGTCATCACCAACCTCTCTTTCAGTAGAGCCTCC 4642
QY 260 TyrThr 261
XX |||||||
Db 4643 TACACA 4648
XX
XX RESULT 11
ACC72052
ID ACC72052 standard; DNA; 9695 BP.
XX
XX AC ACC72052;
XX
XX DT 08-JUL-2003 (first entry)
XX
XX DE BCU0205B gene #SEQ ID 81.
XX
XX KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO2003029421-A2.
XX
XX PD 10-APR-2003.
XX
XX PF 02-OCT-2002; 2002WO-US031287.
XX
XX PR 03-OCT-2001; 2001US-0326526P.
XX 14-MAY-2002; 2002US-00144194.
XX
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX WPI; 2003-381623/36.
XX P-PSDB; ABR58318.
```

XX New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;

Alignment Scores: 8.41e-79 Length: 9695
Pred. No.: 863.50 Matches: 162
Score: 77.10% Conservative: 40
Percent Similarity: 61.83% Mismatches: 59
Best Local Similarity: 62.39% Indels: 1
Query Match: 7 Gaps: 1
DB:

US-10-029-020-14_COPY_1490_1750 (1-261) x ACC72052 (1-9695)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20
Db AAGATTAAACCGTCTACGCCAGGTAAACAACCAACGGGGAGATCTGCCCTTTAGCTGGGCA 4559
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspAspGly 40
Db GCCTCGGACTGCGACTGCAAAACACGATGTCAATTGCAACTGCTATTTCAGGAGATGATGCC 4619
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db TAGCGGACTGATGCCATCTTGAATTCCCATCATCTCTAGCTAGCTCCAGATGGTACC 4679
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProphe 80
Db ATTTACATTGCAGACCTTGGAAATATTGGATCAGGGCGGTGACAGCAAGCAAGCCTGTT 4739
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuphe 100
Db CTTAATGCCCTTCAACACAGTATGAGCTGCATCCCCCGGAGAGCAGGAGTTATATGTTTC 4799
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db AACGCTGATGGCATCCCAATACACTGTGAGCCTGGTGACAGGGGAGTACTTGTACAAT 4859
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db TTCACATATAGTACTGACAAATGATGTCATGAAATTGATTGACAAATATGGGAATTCCTTG 4919
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValProAspGlyGlnVal 160
Db AAGATCCGTCGGGACAGCAGTGGCATGCCCCGTACCTGCTCATGCTGACCAACAGATC 4979
QY 161 TyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db ATCACCCCTCACCGTGGGCACCAATGGAGGCTCAAGTCTGTCCACACAGAACCTGGAG 5039
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuAlaThrLysSerAsnGluAsn 200
Db CTGGTCTCATGACCTATGATGGCAACACTGGGCTCCTGGCCACCAAGAGCGATGAAACA 5099

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db GGATGGACGACTTCTATGACTATGACACGAGCGCGCTGACCAACGTGACGGCCCC 5159
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db ACGGGGTGGTAACCACTCTGCACCGGGAATGGAGAAATCTATTACCATTTGACATTGAG 5219
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db AACTCCAACCGTGATGATGACGTCACTGTCTATCATCAACCACTCTCTTCAGTAGAGCCTCC 5279
QY 260 TyrThr 261
Db TACACA 5285
RESULT 12
AAS14089
ID AAS14089 standard; DNA; 9729 BP.
XX
AC AAS14089;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human FCTR3f DNA sequence.
XX
KW Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greennow type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH 5'UTR 1..209 /*tag= b
FT 210..8384 /*tag= a
FT /product= "Human FCTR3f"
FT 8385..9729
FT /*tag= c
XX
PD WO200166747-A2.
XX
PD 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US0007160.
XX
PR 03-MAR-2000; 2000US-0186592P.
PR 03-MAR-2000; 2000US-0186718P.
PR 06-MAR-2000; 2000US-0187293P.
PR 06-MAR-2000; 2000US-0187294P.
PR 17-MAR-2000; 2000US-0190400P.
PR 07-APR-2000; 2000US-0196018P.
PR 03-JAN-2001; 2001US-0259548P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;

PT New FCTR3X polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTR3X-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.

XX Claim 9; Page 34-37; 155pp; English.

XX The invention relates to FCTR3X polypeptides and the polynucleotides
XX encoding them. The sequences of the invention are useful for the
XX manufacture of a medicament for diagnosing and treating disorders
XX associated with the FCTR3X polypeptide, such as colorectal cancer,
XX adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
XX autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
XX tumours, mammary tumours, human gliomas, astrocytomas, renal cell
XX carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
XX cell and granular cell carcinomas, neurological disorders,
XX neurodegenerative disorders, nerve trauma, familial myelodysplastic
XX syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
XX conditions, immunological disorders, allergy and infection, asthma, lung
XX diseases, male and female reproductive disorders, deafness, glycoprotein
XX deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
XX C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansonii
XX infection, spinocerebellar ataxia, plasmodium falciparum infection,
XX Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
XX represents cDNA encoding an FCTR3X polypeptide of the invention.

SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores: Length: 9729
Pred. No.: 8,45e-79 Matches: 162
Score: 863.50 Conservative: 40
Percent Similarity: 77.10% Mismatches: 59
Best Local Similarity: 61.83% Indels: 1
Query Match: 62.39% Gaps: 1
DB:

US-10-029-020-14_COPY_1490_1750 (1-261) x ADB32028 (1-9729)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluLeuSerLeuValAlaGlyAla 20
Db 4548 AAGATTAAACCGTCTACGCCAGGTAAACAACCAACGGGGAGATCGCTTTTAGCTGGGCA 4607
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4608 GCCTCGGACTGCGACTGCAAAACGATGTCAATTGCAACTGCTATTCAGGAGATGATGCC 4667
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4668 TACGCGACTGATGCCATCTTGAATTCCCATCATCTTGTAGTCCAGATGGTACC 4727
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4728 ATTTACATGTCAGACCTTGGAAATATTCGGATCAGGGCGGTCTACCAAGAACAAAGCCTGTT 4787
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4788 CTTAATGCCTTCAACCCAGTATGAGGCTGCATCCCCCGGAGAGCAGAGTTATATGTTTC 4847
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4848 AACGCTGATGGCATCCCAATACACTGTGAGCTGGTGACAGGGGAGTACTTGTAACAAT 4907
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 4908 TTCACATATAGTACTGACAATGATGTCACTGAATTGATGACAATATGGGAATTCCTCG 4967
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160
Db 4968 AAGATCCGTCGGGACAGCAGTGGCATGCCCGCTGCTCATGCTGCAACCAAGATC 5027
QY 161 TyrTyrValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5028 ATCACCCCTCACCGTGGGACCAATGGAGGCTCAAGATCGTGTCCACACAGAACCTGGAG 5087

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5088 CTTGGTCTCATGACCTATGATGGCAACACTCTGGCTCTCTGGCCACCAAGAGCGATGAAACA 5147
QY 201 GlyTyrThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5148 GGATGGACGACTTCTATGACTATGACCAACGAGGCGCGCTGACCAACGTCGCGCCCC 5207
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5208 ACGGGGTGGTAACCACTCTGCACCGGAAATGGAGAAATCTATTACCATTTGACATTGAG 5267
QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5268 AACTCCAAACCGTATGATGACGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5327
QY 260 TyrThr 261
Db 5328 TACACA 5333
RESULT 14
AAS14085
ID AAS14085 standard; DNA; 9826 BP.
XX
AC AAS14085;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human FCTR3b DNA sequence.
XX
KW Human; FCTR3X; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansonii infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greenouw type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3b; neurestin-like protein.

OS Homo sapiens.

Key Location/Qualifiers
5'UTR 1..279
FT /*tag= b
FT 280..8481
FT /*tag= a
FT /*product= "Human FCTR3b"
FT 8482..9826
FT /*tag= c

WO200166747-A2.

13-SEP-2001.

05-MAR-2001; 2001WO-US007160.

03-MAR-2000; 2000US-0186592P.

03-MAR-2000; 2000US-0186718P.

06-MAR-2000; 2000US-0187293P.

06-MAR-2000; 2000US-0187294P.

17-MAR-2000; 2000US-0190400P.

07-APR-2000; 2000US-0196018P.

03-JAN-2001; 2001US-0259548P.

XX PA (CURA-) CURAGEN CORP.

XX PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;

XX PI Macdougall J, Mishra V, Mezes PS, Rastelli L;

XX DR WPI; 2001-596837/67.

XX DR P-PSDB; AAU08680.

PT Novel polypeptides designated as FCTR3b polypeptides, useful in detection, prevention and treatment of a broad range of pathological states.

PT

PS Claim 9; Page 33-35; 215pp; English.

XX

CC The invention relates to human FCTR3b polypeptides, FCTR1-FCTR7, and the nucleic acids encoding them. These sequences are useful for the treatment or prevention of numerous disorders including myelogenous leukaemia, CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal CC alloimmune thrombocytopaenia, neurological disorders, neurodegenerative CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie CC -tooth neuropathy, demyelinating Gardner syndrome, familial CC myelodysplastic syndrome, mental health conditions, immunological CC disorders, allergy and infection, bronchial asthma, Avellino type CC eosinophilia, lung diseases, reproductive disorders, infertility, male CC and female reproductive disorders, deafness, glycoprotein Ia deficiency, CC desmoid disease, turcotte syndrome, liver cirrhosis, hepatitis C, gastric CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni CC infection, spinocerebellar ataxia, Plasmodium falciparum parasitaemia, CC Corneal dystrophy-Greenou type I, Corneal dystrophy-lattice type I and CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding FCTR3b, a neurestin-like protein

XX

SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.55e-79 Length: 9826

Score: 863.50 Matches: 162

Percent Similarity: 77.10% Conservative: 40

Best Local Similarity: 61.83% Mismatches: 59

Query Match: 62.39% Indels: 1

DB: 5 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x AAS14085 (1-9826)

QY 1 LysileAsnArgIleArgGlnValThrThrSerGlyGluIleSerLeuValAlaGlyAla 20

Db 4645 AAGATTAAACCGTCTACGCCAGGTAAACCAACGGGGAGATCTGCCCTTTAGCTGGGCA 4704

QY 21 ProSerGlyCysAspCysLysAsnAlaAsnAspCysAspCysPheSerGlyAspGly 40

Db 4705 GCCTCGGACTCGGACTGCAAAACGATGTCAATGTCAACTGTCTATTCAGGAGATGTC 4764

QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60

Db 4765 TACGCGACTGATGCCATCTTGAATTCGCCATCATCTTAGCTGTAGTCCAGATGGTACC 4824

QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgPheIleArgLysAsnLysProPhe 80

Db 4825 ATTTACATTCAGACCTTGGAATAATTTCGGATCAGGGCGGTTCAGCAAGAACAAGCCTGTT 4884

QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100

Db 4885 CTTAATGCTTCAACCCAGTATGAGGCTGCATCCCCGGAGAGCAGGAGTTATGTTTTC 4944

QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120

Db 4945 AACGCTGATGGCATCCCAATACTGTGAGCCTGGTGACAGGGGAGTACTTGTACAAAT 5004

QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140

Db 5005 TTCACATATAGTACTGACAATGATGCTACTGAATTTGATGACAATAATGGGAATTCCTG 5064

QY 141 AsnValArgArgAspSerThrGlyMetProLeuTyrLeuValValProAspGlyGlnVal 160

Db 5065 AAGATCCGTCGGGACAGCAGTGGCATGCCCGTCACTGCTCATGCCTGACAAACGATC 5124

QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180

Db 5125 ATCACCCCTCACCGTGGGACCAATGGAGGCTCAAAAGTCGTGTCCACACAGAACCTGGAG 5184

QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200

Db 5185 CTTGGTCTCATGACCTATGATGGCAACACTGGGCTCCTGGCCACCAAGAGCGATGAACA 5244

QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220

Db 5245 GGATGGACGACTTCTATGACTATGACCAACGAGGCGCTGACCAACGTGACGCGCCCC 5304

QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240

Db 5305 ACGGGGTGGTAACCACTGTGCACCGGAAATGGAGAAATCTATTACCATTGACATGAG 5364

QY 241 ThrSerSerLys---AspAspValThrIleThrThrAsnLeuSerAlaSerGlyAlaPhe 259

Db 5365 AACTCCAACCGTGATGATGACGTCACTGTCTATCAACCACTCTCTTCTAGTAGGCTCC 5424

QY 260 TyrThr 261

Db 5425 TACACA 5430

RESULT 15

ADB32023

ID ADB32023 standard; cDNA; 9826 BP.

XX

AC ADB32023;

XX

DT 04-DEC-2003 (first entry)

XX

DE Human FCTR3b cDNA.

XX

KW Human; gene; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli; myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia; ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma; renal cell carcinoma; melanoma; clear cell carcinoma; granular cell carcinoma; neurological disorder; neurodegenerative disorder; nerve trauma; familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy; Gardner syndrome; mental health condition; immunological disorder; allergy; asthma; lung disease; reproductive disorder; deafness; glycoprotein deficiency; desmoid tumour; turcotte syndrome; liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease; diabetes; schistosoma mansoni infection; spinocerebellar ataxia; plasmodium falciparum infection; Groenou's corneal dystrophy; lattice corneal dystrophy.

XX

OS Homo sapiens.

XX

PN US2003087816-A1.

XX

PD 08-MAY-2003.

XX

PF 05-MAR-2001; 2001US-00800198.

XX

PR 03-MAR-2000; 2000US-0186592P.

XX

PA (VERM/) VERMET C.

PA (FERN/) FERNANDES E.

PA (SHIM/) SHIMKETS R.

PA (HERR/) HERRMANN J.

PA (MAJU/) MAJUMDER K.

PA (MACD/) MACDOUGALL J.

PA (MISH/) MISHRA V.

PA (MEZE/) MEZES P S.

PA (RAST/) RASTELLI L.

XX

PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;

PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2003-625633/59.
DR P-PSDB; ADB32024.

XX New FCTR peptide and encoding polynucleotide, useful for preventing
PT or treating FCTR-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.

XX Claim 9; Page 29-32; 155pp; English.

XX The invention relates to FCTR peptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTR polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTR polypeptide of the invention.

XX
SQ Sequence 9826 BP; 2567 A; 2645 C; 2476 G; 2138 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8.55e-79 Length: 9826
Score: 863.50 Matches: 162
Percent Similarity: 77.10% Conservative: 40
Best Local Similarity: 61.83% Mismatches: 59
Query Match: 62.39% Indels: 1
DB: 9 Gaps: 1

US-10-029-020-14_COPY_1490_1750 (1-261) x ADB32023 (1-9826)

QY 1 LysIleAsnArgIleArgGlnValThrThrSerGlyGluLeuSerLeuValAlaGlyAla 20
Db 4645 AAGATTAAACCGTCTACGCAGGTAAACAACCAACGGGAGATCGCTTTTAGCTGGGGCA 4704
QY 21 ProSerGlyCysAspCysLysAsnAspAlaAsnCysAspCysPheSerGlyAspGly 40
Db 4705 GCCTCGGACTGCGACTGCAAAAACGATGTCAATTGCAACTGCTATTTCAGGAGATGATGCC 4764
QY 41 TyrAlaLysAspAlaLysLeuAsnThrProSerSerLeuAlaValCysAlaAspGlyGlu 60
Db 4765 TAGCGACTGATGCCATCTTGAATTCCTCATCTAGCTAGCTCCAGATGCTACC 4824
QY 61 LeuTyrValAlaAspLeuGlyAsnIleArgIleArgPheIleArgLysAsnLysProPhe 80
Db 4825 ATTTACATTGCAGACCTTGGAAATATTTCGGATCAGGGCGGTTCAGCAAGAACAGCCTGTT 4884
QY 81 LeuAsnThrGlnAsnMetTyrGluLeuSerSerProIleAspGlnGluLeuTyrLeuPhe 100
Db 4885 CTTAATGCCTTCAACCAAGTATGAGGCTGCATCCCCCGGAGAGCAGGATTATATGTTTC 4944
QY 101 AspThrThrGlyLysHisLeuTyrThrGlnSerLeuProThrGlyAspTyrLeuTyrAsn 120
Db 4945 AACGCTGATGGCATCCACCAATACACTGTGAGCCTGGTGACAGGGGAGTACTTGTACAAAT 5004
QY 121 PheThrTyrThrGlyAspGlyAspIleThrLeuIleThrAspAsnAsnGlyAsnMetVal 140
Db 5005 TTCACATATAGTACTGACAAATGATGTCACACTGAATTGATTGACAATAATGGGAATTCCTCG 5064
QY 141 AsnValArgArgAspSerThrGlyMetProLeuTrpLeuValValProAspGlyGlnVal 160

Db 5065 AAGATCCGTCGGGACAGCAGTGGCATGCCCGTCACCTGCTCATGCTGACAACCATC 5124
QY 161 TyrTrpValThrMetGlyThrAsnSerAlaLeuLysSerValThrThrGlnGlyHisGlu 180
Db 5125 ATCACCTCACCGTGGGCACCAATGGAGGCCTCAAAGTCGTGTCACACACAGACCTGGAG 5184
QY 181 LeuAlaMetMetThrTyrHisGlyAsnSerGlyLeuLeuAlaThrLysSerAsnGluAsn 200
Db 5185 CTTGTCTCATGACCTATGATGGCAACACTGGGCTCCTGGCCACCAAGAGCGATGAAACA 5244
QY 201 GlyTrpThrThrPheTyrGluTyrAspSerPheGlyArgLeuThrAsnValThrPhePro 220
Db 5245 GGATGGACGACTTCTATGACTATGACCAACGAGCCCGCTGACCAACGTGACGCGCCCC 5304
QY 221 ThrGlyGlnValSerSerPheArgSerAspThrAspSerSerValHisValGlnValGlu 240
Db 5305 ACGGGGTGGTAACCACTCTGCACCGGGAATGGAGAAATCTATTACCATTGACATTGAG 5364
QY 241 ThrSerSerLys---AspAspValThrIleThrThrThrAsnLeuSerAlaSerGlyAlaPhe 259
Db 5365 AACTCCAACCGTGATGATGACGTCACTCATCACCAACCTCTCTTCAGTAGAGGCCTCC 5424
QY 260 TyrThr 261
Db 5425 TACACA 5430

Search completed: August 14, 2004, 02:31:38
Job time : 514.726 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 302.635 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400
Perfect score: 782
Sequence: 1 YIRRIFFSGNVTNILELRNK.....LLGSNDLTSARPLSCDSVMD 151

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10029020/runat_06082004_112218_29331/app query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10029020@cgn_1_1_2156@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA :
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
§				

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

1	782	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
2	782	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
3	759.5	97.1	8355	13	US-10-383-201-55	Sequence 55, Appl
4	759.5	97.1	8438	13	US-10-042-865-1	Sequence 1, Appli
5	625	79.9	8409	9	US-09-808-602-79	Sequence 79, Appl
6	625	79.9	8409	10	US-09-800-198-67	Sequence 67, Appl
7	619	79.2	8473	17	US-10-038-854-39	Sequence 39, Appl
8	619	79.2	8487	17	US-10-038-854-41	Sequence 41, Appl
9	619	79.2	8645	17	US-10-038-854-37	Sequence 37, Appl
10	619	79.2	8675	17	US-10-038-854-35	Sequence 35, Appl
11	613	78.4	6560	9	US-09-808-602-76	Sequence 76, Appl
12	613	78.4	6560	10	US-09-800-198-64	Sequence 64, Appl
13	613	78.4	9058	16	US-10-144-194A-79	Sequence 79, Appl
14	613	78.4	9695	16	US-10-144-194A-81	Sequence 81, Appl
15	613	78.4	9729	9	US-09-808-602-12	Sequence 12, Appl
16	613	78.4	9729	10	US-09-800-198-12	Sequence 12, Appl
17	613	78.4	9826	9	US-09-808-602-7	Sequence 7, Appli
18	613	78.4	9826	10	US-09-800-198-7	Sequence 7, Appli
19	610	78.0	8689	9	US-09-808-602-78	Sequence 78, Appl
20	610	78.0	8689	10	US-09-800-198-66	Sequence 66, Appl
21	610	78.0	8797	9	US-09-808-602-74	Sequence 74, Appl
22	610	78.0	8797	9	US-09-808-602-77	Sequence 77, Appl
23	610	78.0	8797	10	US-09-800-198-62	Sequence 62, Appl
24	610	78.0	8797	10	US-09-800-198-65	Sequence 65, Appl
25	602.5	77.0	8575	13	US-10-072-012-143	Sequence 143, App
26	534.5	68.4	12880	16	US-10-295-027-927	Sequence 927, App
27	413	52.8	247	15	US-10-029-386-16574	Sequence 16574, A
28	413	52.8	557	15	US-10-029-386-2874	Sequence 2874, Ap
29	211	27.0	751	13	US-10-027-632-160821	Sequence 160821,
30	211	27.0	751	16	US-10-027-632-160821	Sequence 160821,
31	111	14.2	789	13	US-10-424-599-25144	Sequence 25144, A
32	101	12.9	1482	9	US-09-938-842A-1949	Sequence 1949, Ap
33	101	12.9	1482	11	US-09-938-842A-1949	Sequence 1949, Ap
34	98	12.5	1480	13	US-10-424-599-30579	Sequence 30579, A
35	98	12.5	2276	13	US-10-424-599-30577	Sequence 30577, A
36	93	11.9	1719	13	US-10-425-114-20338	Sequence 20338, A
37	89	11.4	1851	17	US-10-437-963-88447	Sequence 88447, A
38	88.5	11.3	1431	13	US-10-424-599-132367	Sequence 132367,
39	88.5	11.3	1925	17	US-10-437-963-11176	Sequence 11176, A
40	85	10.9	2550	16	US-10-104-047-72	Sequence 72, Appl
41	83	10.6	1735	13	US-10-425-114-707	Sequence 707, App
42	82.5	10.5	2061	13	US-10-424-599-13785	Sequence 13785, A
43	82	10.5	4377	9	US-09-815-242-9570	Sequence 9570, Ap
44	81.5	10.4	450	15	US-10-029-386-16526	Sequence 16526, A
45	81.5	10.4	555	15	US-10-029-386-2826	Sequence 2826, Ap

ALIGNMENTS

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 9.15e-108 Length: 8354
Score: 782.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-383-201-43 (1-8354)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTAGAGGATCTTCCCTCTGGAATGTCAACAATCTTAGAGCTGAGGAATAAA 3841
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3842 GATTTTCAGACATAGTCACAGTCCAGCACACAATACTACCTGGCCACAGACCCCATGAGT 3901
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGGCCGTCTTCCCTTTCTGACAGCAACACAGCCGGCGGGTCTTTAAATCAAGTCCACTGTG 3961
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCCAGTGCCTC 4021
QY 81 PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCCACACTACCAAT 4081
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCCAGGGGTATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGATC 4141
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCACATCA 4201
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTCAGCTGTGATTCTGTCATGGAT 4234

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US2004003971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590

; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.: 9.15e-108 Length: 8354
Score: 782.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-029-020-13 (1-8354)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTAGAGGATCTTCCCTCTGGAATGTCAACAATCTTAGAGCTGAGGAATAAA 3841
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3842 GATTTTCAGACATAGTCACAGTCCAGCACACAATACTACCTGGCCACAGACCCCATGAGT 3901
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGGCCGTCTTCCCTTTCTGACAGCAACACAGCCGGCGGGTCTTTAAATCAAGTCCACTGTG 3961
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCCAGTGCCTC 4021
QY 81 PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCCACACTACCAAT 4081
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCCAGGGGTATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGATC 4141
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCACATCA 4201
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTCAGCTGTGATTCTGTCATGGAT 4234

RESULT 3
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO
; FILE REFERENCE: 21402-568A

; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 2.6e-104 Length: 8355
Score: 759.50 Matches: 151
Percent Similarity: 96.79% Conservative: 0
Best Local Similarity: 96.79% Mismatches: 0
Query Match: 97.12% Indels: 5
DB: 13 Gaps: 2

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-383-201-55 (1-8355)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeu-----Arg 18
3754 TACATTAGAGGATCTTCCCCTCTGGAAATGTCAACAAATCCTAGAGCTGAGGGTCAGA 3813
QY 19 AsnLysAspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPro 38
3814 AATAAAGATTTCAGACATAGTACAGTCCAGCACACAAATACTACCTGGCCACAGACCCC 3873
QY 39 MetSerGlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSer 58
3874 ATGAGTGGGGCCGCTTCTCCTTCTGACAGCAACAGCCGGGGTCTTTTAAATCAAGTCC 3933
QY 59 ThrValValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGln 78
3934 ACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCAG 3993
QY 79 CysLeuProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeu 98
3994 TGCCTCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCCACACTC 4053
QY 99 ThrAsnProArg-----GlyIleThrValAspLysPheGlyLeuIleTyrPheVal 115
4054 ACCAATCCAGGGGTCCCCCAGGCATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTG 4113
QY 116 AspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySer 135
4114 GATGGCACCATGATCAGACGCATCGATCAGAATGGGATCATCTCCACCCCTGCTCGGCTCT 4173
QY 136 AsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMetAsp 151

Db 4174 AATGATCTCACATCAGCCCGGCACTAGCTGTGATTCTGTCTATGGAT 4221
RESULT 4
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennnda
; APPLICANT: Gunther, Brik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:
Pred. No.: 2.64e-104 Length: 8438
Score: 759.50 Matches: 151
Percent Similarity: 96.79% Conservative: 0
Best Local Similarity: 96.79% Mismatches: 0
Query Match: 97.12% Indels: 5
DB: 13 Gaps: 2

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-042-865-1 (1-8438)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeu-----Arg 18
3764 TACATTAGAGGATCTTCCCCTCTGGAAATGTCAACAAATCTAGAGCTGAGGGTCAGA 3823
Db

QY	19	AsnLysAspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPro	38
DB	3824	AATAAAGATTTCAGACATAGTCACAGTCCAGCACAAATACTACCTGGCCACAGACCCC	3883
QY	39	MetSerGlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSer	58
DB	3884	ATGAGTGGGGCCGTCCTTCTTGACAGCACAGCCGGGGGTCTTTAAAAATCAAGTCC	3943
QY	59	ThrValValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGln	78
DB	3944	ACTGTGTTGGTGAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCAG	4003
QY	79	CysLeuProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeu	98
DB	4004	TGCCTCCCCCTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTC	4063
QY	99	ThrAsnProArg-----GlyIleThrValAspLysPheGlyLeuIleTyrPheVal	115
DB	4064	ACCAATCCCAGGGGTCCCCCAGGCATTACAGTGGACAAAGTTTGGCTGATCTACTTCGTG	4123
QY	116	AspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySer	135
DB	4124	GATGGCACCATGATCAGACGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCT	4183
QY	136	AsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMetAsp	151
DB	4184	AATGATCTCACATCAGCCCGGCACCTCAGCTGTGATTCTGTCTATGGAT	4231

```

RESULT 5
US-09-808-602-79
; Sequence 79, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-808-602-79

Alignment Scores:
Pred. No.: 1.14e-83 Length: 8409
Score: 625.00 Matches: 113
Percent Similarity: 90.07% Conservative: 23
Best Local Similarity: 74.83% Mismatches: 15
Query Match: 79.92% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-808-602-79 (1-8409)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3856 TACATTCGGGTATCTTCCCATCCAGGAATGTGACTAGCATATTGGAGCTGAGAAATAAA 3915
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrIleuAlaThrAspProMetSer 40

```

Db	3916	GAGTTTAAACATAGCAACAATCCTGTCTCACAATACTATCTGGCCGTGGACCCCGTTTCG	3977
QY	41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60
Db	3976	GGCTCCCTGTAGTATCAGACACCAACAGCCGACGGATATACAAAGTCAAATCTCTTACT	4035
QY	61	ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu	80
Db	4036	GGCACGAAAGACCTGGCTGGTAATCTGAAGTGGTAGCGGGGACTGGAGAGCAATGCCTG	4095
QY	81	ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn	100
Db	4096	CCCTTTGATGAAGCCAGATGTGGAGATGGAGGGAAAGCAGTGGACGCAACCTAATGAGT	4155
QY	101	ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle	120
Db	4156	CCTCGAGGAATTGCAGTGGATAAGTATGGACTCATGTATTTTGTGATGCCACTATGATT	4215
QY	121	ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer	140
Db	4216	CGAAAAGTGGATCAGATGGAATTATATCAACTCTGCTGGGCTCCAATGACCTAACTGCC	4275
QY	141	AlaArgProLeuSerCysAspSerValMetAsp	151
Db	4276	GTCCGACCTCTAAGCTGTGATTCAGCATGGAT	4308

RESULT 6

US-09-800-198-67

; Sequence 67, Application US/09800198

; Publication No. US20030087816A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrmann, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 67

; LENGTH: 8409

; TYPE: DNA

; ORGANISM: Gallus gallus

US-09-800-198-67

Alignment Scores:					
Pred.	No. :	Length:	8409		
Score:	625.00	Matches:	113		
Percent Similarity:	90.07%	Conservative:	23		
Best local Similarity:	74.83%	Mismatches:	15		
Query Match:	79.92%	Indels:	0		
DB:	10	Gaps:	0		
 US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-800-198-67 (1-8409)					
QY	1	TyrIleArgGilePheProSerGlyAsnValThrAsnIlLeuGlueUargAsnLys	20		
Dd	3856	TACATTGGCGTATCTTC CATCCAGGAATGTGACTAGCATATTGGAGCTGAGAAATAA	391		
QY	21	AspPheArgHisSerHISerProAlaHisLysTy r TyrLeuAlaThraSpPrometSer	40		
Dd	3916	GAGTTTAACA TAGCAACA ATCCTGCTCACAAA TACTAT CTGGCCG TG GACC CG GTT TC G	397		
QY	41	GlyAlavalpheUs er Asp Ser As n Ser Arg Val Phel y s l e L ys Ser Thr Val	60		

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 3976 GGCTCCCTGTACGTATCAGACACCAACAGCCGACGGATATACAAAGTCAAATCTCTTACT 4035
QY 61 ValVallysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4036 GGCACGAAGACCTGGCTGGTAATCTGAAGTGGTAGCGGGACTGGAGAGCAATGCCTG 4095
QY 81 PropheAspAspThrArgCysGlyAspGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4096 CCCTTTGATGAAGCCAGATGTGGAGATGGAGGGAAAGACAGTGGACGCAACCCTAATGAGT 4155
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4156 CCTCGAGGAATTGCAGTGGATAAGTATGGACTCATGTATTTTGTGATGCCACTATGATT 4215
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4216 CGAAAAGTGGATCAGAATGGAATTATATCAACTCTGCTGGGCTCCAATGACCTAACTGCC 4275
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4276 GTCCGACCTCTAAGCTGTGATTCAGCATGGAT 4308

RESULT 7

US-10-038-854-39

; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889

; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:

Pred. No.:	9.65e-83	Length:	8473
Score:	619.00	Matches:	111
Percent Similarity:	90.67%	Conservative:	25
Best Local Similarity:	74.00%	Mismatches:	14
Query Match:	79.16%	Indels:	0
DB:	17	Gaps:	0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-39 (1-8473)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3600 TACGTGCGCGGATATTCCTTCTTGAAATGTAAACAAGTGTCTTAGAACCTAAGAAATAAA 3659
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPrometSer 40
Db 3660 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 3719
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3720 GGAGATCTGTACGTTTCTGACACAAACACCCCGCAGAAATTTATCGCCCAAGTCACTTACG 3779
QY 61 ValVallysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3780 GGGGCAAAAGACTTGACTAAAAATGCAGAAAGTCGTGCGAGGGACAGGGGAGCAATGCCTT 3839
QY 81 PropheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 3840 CCGTTTGACGAGCGAGATGTGGGATGGAGGGAAGGCCGTGTGGAAGCCACACTCATGAGT 3899
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 3900 CCCAAAGGAATGGCAGTTTGATAAGAAATGGATTAATCTACTTTGTTGATGGAACCATGATT 3959
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3960 AGGAAAGTTGACCACAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 4019
QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4020 GCCAGACCTTTAACTTGTGACACCCAGCATG 4049

RESULT 8

US-10-038-854-41

; Sequence 41, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh

```

; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.: 9.68e-83 Length: 8487
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-41 (1-8487)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnIlys 20
Db 3596 TACGTGGCGGATATCCCTTCTGGAAATGTAACAAGTGTCTTAGAACTAAGAAATAAA 3655

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3656 GATTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 3715

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3716 GGAGATCTGTACGTTTCTGACACAAACACCCGAGATTATCGCCCAAGTCACCTACG 3775

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3776 GGGGCAAAAGACTTGACTAAATAATGCAGAAGTCGTGCGAGGACAGGGGAGCAATGCCTT 3835

QY 81 PropheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 3836 CCGTTTGACGAGCGGAGATGTGGGGATGGAGGGAAGCCGTTGGAGCCACACTCATGAGT 3895

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 3896 CCCAAAGGAATGGCAGTTGATAAGATGGATTAACTACTCTTTGTTGATGGAAACCATGATT 3955

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3956 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTGACTTCA 4015

QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4016 GCCAGACCTTTAACTTGTGACACCAGCATG 4045

RESULT 9
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
```


; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.: 9.97e-83 Length: 8645
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-37 (1-8645)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
|||:::|||||
Db 3772 TACGTGCGCGGATATTCCTTCTGGAAATGTAACAAGTGTCTTAGAACTAAGAAATAAA 3831
|||:::|||||
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
|||:::|||||
Db 3832 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 3891
|||:::|||||
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
|||:::|||||
Db 3892 GGAGATCTGTACGTTTCTGCACAAACACCCCGCAGAATTTATCGCCCAAGTCACCTACG 3951
|||:::|||||
QY 61 ValVallLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
|||:::|||||
Db 3952 GGGGCAAAAGACTTGACTAAAATGCAGAAATCGTCGACGGGACAGGGGAGCAATGCCTT 4011
|||:::|||||
QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
|||:::|||||
Db 4012 CCGTTTGACGAGCGAGATGTGGGGATGGAGGAGGCCGCGTGAAGCCACACTCATGAGT 4071
|||:::|||||
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
|||:::|||||
Db 4072 CCCAAGGAATGGCAGTTGATAAGAATGGATTAACTACTTTTGTGATGGAACCATGATT 4131
|||:::|||||
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
|||:::|||||
Db 4132 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 4191
|||:::|||||
QY 141 AlaArgProLeuSerCysAspSerValMet 150
|||:::|||||
Db 4192 GCCAGACCTTTAACTTGTGACACCAGCATG 4221
|||:::|||||

RESULT 10
US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:

; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:
Pred. No.: 1e-82 Length: 8675
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-038-854-35 (1-8675)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
|||:::|||||
Db 3784 TACGTGCGCGGATATTCCTTCTGGAAATGTAACAAGTGTCTTAGAACTAAGAAATAAA 3843
|||:::|||||
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
|||:::|||||
Db 3844 GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 3903
|||:::|||||
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
|||:::|||||
Db 3904 GGAGATCTGTACGTTTCTGCACAAACACCCCGCAGAATTTATCGCCCAAGTCACCTACG 3963
|||:::|||||
QY 61 ValVallLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
|||:::|||||
Db 3964 GGGGCAAAAGACTTGACTAAAATGCAGAAATCGTCGACGGGACAGGGGAGCAATGCCTT 4023
|||:::|||||
QY 81 ProPheAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
|||:::|||||
Db 4024 CCGTTTGACGAGCGAGATGTGGGGATGGAGGAGGCCGCGTGAAGCCACACTCATGAGT 4083
|||:::|||||

```
QY      101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||:::|||||
Db      4084 CCCAAAGGAATGGCAGTTGATAAGAATGGATTAATCTACTTTGTTGATGGAACCATGATT 4143
      |||:::|||||
QY      121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      |||:::|||||
Db      4144 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCA 4203
      |||:::|||||
QY      141 AlaArgProLeuSerCysAspSerValMet 150
      |||:::|||||
Db      4204 GCCAGACCTTTAACTTGTGACACCAGCATG 4233
      |||:::|||||

RESULT 11
US-09-808-602-76
; Sequence 76, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herriman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 6560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-808-602-76

Alignment Scores:
Pred. No.:      5.34e-82      Length:      6560
Score:          613.00      Matches:     112
Percent Similarity: 89.40%      Conservative: 23
Best Local Similarity: 74.17%      Mismatches:  16
Query Match:    78.39%      Indels:      0
DB:              9      Gaps:          0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-808-602-76 (1-6560)

QY      1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
      |||:::|||||
Db      662 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA 721
      |||:::|||||
QY      21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
      |||:::|||||
Db      722 GAGTTTAAACATAGCAACAACCCAGCACACACAGGAGAATCTACCGCGTCAAGTCTCTGAGT 781
      |||:::|||||
QY      41 GlyAlaValPheLeuSerAspSerAsnSerGluValAlaGlyThrGlyAspGlnCysLeu 80
      |||:::|||||
Db      782 GGCTCGCTCTACGTGCCGACACCAACAGCAGGAGAATCTACCGCGTCAAGTCTCTGAGT 841
      |||:::|||||
QY      61 ValVallysAspLeuValLysAsnSerGluValAlaGlyThrGlyAspGlnCysLeu 80
      |||:::|||||
Db      842 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGGACGGGAGACGAGTGCTCTA 901
      |||:::|||||
QY      81 PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
      |||:::|||||
Db      902 CCCTTTGATGAAGCCCGCTGCGGGATGGAGGAAGGCCATAGATGCAACCCCTGATGAGC 961
      |||:::|||||
QY      101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||:::|||||
Db      902 CCCTTTGATGAAGCCCGCTGCGGGATGGAGGAAGGCCATAGATGCAACCCCTGATGAGC 961
      |||:::|||||
QY      101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||:::|||||
```

```
Db      962 CCGAGAGGTATTGCAGTAGACAAGAATGGGCTCATGTACTTTGTGATGCCACCATGATC 1021
      |||:::|||||
QY      121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      |||:::|||||
Db      1022 CGGAAGTTTGACCAGAAATGGAATCATCTCCACCCCTGCTGGGCTCCAATGACCTCACTGCC 1081
      |||:::|||||
QY      141 AlaArgProLeuSerCysAspSerValMetAsp 151
      |||:::|||||
Db      1082 GTCCGGCCGCTGAGCTGTGATTCCAGCATGGAT 1114
      |||:::|||||

RESULT 12
US-09-800-198-64
; Sequence 64, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 6560
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-800-198-64

Alignment Scores:
Pred. No.:      5.34e-82      Length:      6560
Score:          613.00      Matches:     112
Percent Similarity: 89.40%      Conservative: 23
Best Local Similarity: 74.17%      Mismatches:  16
Query Match:    78.39%      Indels:      0
DB:              10      Gaps:          0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-800-198-64 (1-6560)

QY      1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
      |||:::|||||
Db      662 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA 721
      |||:::|||||
QY      21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
      |||:::|||||
Db      722 GAGTTTAAACATAGCAACAACCCAGCACACACAGGAGAATCTACCGCGTCAAGTCTCTGAGT 781
      |||:::|||||
QY      41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
      |||:::|||||
Db      782 GGCTCGCTCTACGTGCCGACACCAACAGCAGGAGAATCTACCGCGTCAAGTCTCTGAGT 841
      |||:::|||||
QY      61 ValVallysAspLeuValLysAsnSerGluValAlaGlyThrGlyAspGlnCysLeu 80
      |||:::|||||
Db      842 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGGACGGGAGACGAGTGCTCTA 901
      |||:::|||||
QY      81 PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
      |||:::|||||
Db      902 CCCTTTGATGAAGCCCGCTGCGGGATGGAGGAAGGCCATAGATGCAACCCCTGATGAGC 961
      |||:::|||||
QY      101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||:::|||||
Db      962 CCGAGAGGTATTGCAGTAGACAAGAATGGGCTCATGTACTTTGTGATGCCACCATGATC 1021
      |||:::|||||
QY      121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      |||:::|||||
```

```
Db      1022 CGGAAGGTTGACCAGAATGGAATCATCTCCACCCTGCTGGGCTCCAATGACCTCACTGCC 1081
QY      141 AlaArgProLeuSerCysAspSerValMetAsp 151
      |||||||
Db      1082 GTCCGGCCGCTGAGCTGTGATTCCAGCATGGAT 1114

RESULT 13
US-10-144-194A-79
; Sequence 79, Application US/10144194A
; Publication No. US20030215809A1
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79
; LENGTH: 9058
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (44)..(7699)
US-10-144-194A-79

Alignment Scores:
Pred. No.:      8.95e-82      Length:      9058
Score:          613.00      Matches:      112
Percent Similarity: 89.40%      Conservative: 23
Best Local Similarity: 74.17%      Mismatches: 16
Query Match:    78.39%      Indels:      0
DB:            16      Gaps:      0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-144-194A-79 (1-9058)

QY      1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
      |||||||
Db      3146 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA 3205

QY      21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
      :::::
Db      3206 GAGTTTAAACATAGCAACAACCCAGCACACAAGTACTACTTGGCAGTGACCCCGTGTCC 3265

QY      41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
      ||:::
Db      3266 GGCTCGCTCTACGTGTCGACACCAACAGCAGGAGAATCTACCCGCTCAAGTCTCTGAGT 3325

QY      61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
      |||||||
Db      3326 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGGACGGGAGAGCAGTGCTA 3385

QY      81 ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
      ||:::
Db      3386 CCCTTTGATGAAGCCCGCTGCGGGGATGGAGGAAGGCCATAGATGCAACCCTGATGAGC 3445

QY      101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||||||
Db      3446 CCGAGAGGTATTGCAGTAGACAAGAATGGGCTCATGTACTTTGTGATGCCACCATGATC 3505

QY      121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      ||:::
Db      3506 CGGAAGGTTGACCAGAATGGAATCATCTCCACCCTGCTGGGCTCCAATGACCTCACTGCC 3565

QY      141 AlaArgProLeuSerCysAspSerValMetAsp 151
      |||||||
Db      3566 GTCCGGCCGCTGAGCTGTGATTCCAGCATGGAT 3598

RESULT 14
US-10-144-194A-81
; Sequence 81, Application US/10144194A
; Publication No. US20030215809A1
```

```
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies Inc
; TITLE OF INVENTION: Regulated Breast Cancer Genes
; FILE REFERENCE: 3U 103 R1
; CURRENT APPLICATION NUMBER: US/10/144,194A
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81
; LENGTH: 9695
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (435)..(8336)
US-10-144-194A-81

Alignment Scores:
Pred. No.:      9.98e-82      Length:      9695
Score:          613.00      Matches:      112
Percent Similarity: 89.40%      Conservative: 23
Best Local Similarity: 74.17%      Mismatches: 16
Query Match:    78.39%      Indels:      0
DB:            16      Gaps:      0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-10-144-194A-81 (1-9695)

QY      1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
      |||||||
Db      3783 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA 3842

QY      21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
      :::::
Db      3843 GAGTTTAAACATAGCAACAACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTGTCC 3902

QY      41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
      ||:::
Db      3903 GGCTCGCTCTACGTGTCGACACCAACAGCAGGAGAATCTACCCGCTCAAGTCTCTGAGT 3962

QY      61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
      |||||||
Db      3963 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGGACGGGAGAGCAGTGCTA 4022

QY      81 ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
      ||:::
Db      4023 CCCTTTGATGAAGCCCGCTGCGGGGATGGAGGAAGGCCATAGATGCAACCCTGATGAGC 4082

QY      101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
      |||||||
Db      4083 CCGAGAGGTATTGCAGTAGACAAGAATGGGCTCATGTACTTTGTGATGCCACCATGATC 4142

QY      121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
      ||:::
Db      4143 CGGAAGGTTGACCAGAATGGAATCATCTCCACCCTGCTGGGCTCCAATGACCTCACTGCC 4202

QY      141 AlaArgProLeuSerCysAspSerValMetAsp 151
      |||||||
Db      4203 GTCCGGCCGCTGAGCTGTGATTCCAGCATGGAT 4235

RESULT 15
US-09-808-602-12
; Sequence 12, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
```


FILE REFERENCE: 15966-697 CIP
CURRENT APPLICATION NUMBER: US/09/808,602
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/800,198
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 9729
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (210)..(8381)
US-09-808-602-12

Alignment Scores:
Pred. No.: 1e-81 Length: 9729
Score: 613.00 Matches: 112
Percent Similarity: 89.40% Conservative: 23
Best Local Similarity: 74.17% Mismatches: 16
Query Match: 78.39% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-808-602-12 (1-9729)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3831 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA 3890

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3891 GAGTTTAAACATAGCAACAACCCAGCACACAGTACTACTTGGCAGTGGACCCCGTGTC 3950

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3951 GGCTCGCTCTACGTGTCCGACACCAACAGCAGGAGAATCTACCGCGTCAAGTCTCTGAGT 4010

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4011 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGGACGGGAGAGCAGTGTCTA 4070

QY 81 PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4071 CCCTTTGATGAAGCCCGCTGCGGGGATGGAGGAAGGCCATAGATGCAACCCCTGATGAGC 4130

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4131 CCGAGAGGTATTGCAGTAGACAAGAATGGGCTCATGTACTTTGTGATGCCACCATGATC 4190

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4191 CGGAAGGTTGACCAGAATGGAATCATCTCCACCCCTGCTGGGCTCCAATGACCTCACTGCC 4250

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4251 GTCCGGCCGCTGAGCTGTGATTCAGCATGGAT 4283

Search completed: August 14, 2004, 19:18:15
Job time : 333.635 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 44.6731 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400 ;
Perfect score: 782
Sequence: 1 YIRIFPSGNVTNILELRNK.....LLGSNDLTSARPLSCDSVMD 151

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1.1.258@runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	90.5	11.6	4403765	3	US-09-103-840A-2
C 2	90.5	11.6	4411529	3	US-09-103-840A-1
3	83	10.6	4403765	3	US-09-103-840A-2
4	83	10.6	4411529	3	US-09-103-840A-1
5	81	10.4	9607	4	US-08-961-527-80
6	80	10.2	978	4	US-09-134-000C-790
C 7	79.5	10.2	2736	4	US-09-235-451-3
8	78.5	10.0	14042	3	US-08-652-877-85
9	78.5	10.0	14044	3	US-08-652-877-89
10	78.5	10.0	14080	3	US-08-652-877-87
11	78.5	10.0	14083	3	US-08-476-515A-83
12	78.5	10.0	14086	3	US-08-652-877-83

C 13	76.5	9.8	6943	4	US-09-453-702B-213	Sequence 213, Appl
14	76	9.7	2073	4	US-09-134-001C-2101	Sequence 2101, Ap
C 15	75	9.6	2797	4	US-09-453-702B-244	Sequence 244, App
16	74.5	9.5	1302	4	US-09-252-991A-14317	Sequence 14317, A
C 17	74.5	9.5	1479	4	US-09-252-991A-14286	Sequence 14286, A
C 18	73	9.3	2028	2	US-08-933-750C-73	Sequence 73, Appl
C 19	73	9.3	2028	3	US-09-234-613-73	Sequence 73, Appl
20	73	9.3	4351	4	US-09-060-299-24	Sequence 24, Appl
21	73	9.3	4351	4	US-09-402-923A-24	Sequence 24, Appl
22	73	9.3	4843	4	US-09-060-299-2	Sequence 2, Appli
23	73	9.3	4843	4	US-09-402-923A-2	Sequence 2, Appli
24	73	9.3	4915	4	US-09-060-299-28	Sequence 28, Appl
25	73	9.3	4915	4	US-09-402-923A-28	Sequence 28, Appl
26	73	9.3	4959	4	US-09-060-299-38	Sequence 38, Appl
27	73	9.3	4959	4	US-09-402-923A-38	Sequence 38, Appl
28	73	9.3	5022	4	US-09-060-299-32	Sequence 32, Appl
29	73	9.3	5022	4	US-09-402-923A-32	Sequence 32, Appl
30	73	9.3	5098	4	US-09-060-299-1	Sequence 1, Appli
31	73	9.3	5098	4	US-09-402-923A-1	Sequence 1, Appli
32	73	9.3	5125	4	US-09-060-299-26	Sequence 26, Appl
33	73	9.3	5125	4	US-09-402-923A-26	Sequence 26, Appl
34	73	9.3	5162	4	US-09-060-299-33	Sequence 33, Appl
35	73	9.3	5162	4	US-09-402-923A-33	Sequence 33, Appl
36	73	9.3	5166	4	US-09-060-299-23	Sequence 23, Appl
37	73	9.3	5166	4	US-09-402-923A-23	Sequence 23, Appl
38	73	9.3	5263	4	US-09-060-299-31	Sequence 31, Appl
39	73	9.3	5263	4	US-09-402-923A-31	Sequence 31, Appl
40	72.5	9.3	885	4	US-09-446-301A-3	Sequence 3, Appli
41	72.5	9.3	885	4	US-09-099-932-3	Sequence 3, Appli
42	72.5	9.3	1607	4	US-09-446-301A-16	Sequence 16, Appl
43	72.5	9.3	1607	4	US-09-099-932-50	Sequence 50, Appl
44	72	9.2	1599	4	US-09-060-299-37	Sequence 37, Appl
45	72	9.2	1599	4	US-09-402-923A-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g

Alignment Scores:
Pred. No.: 1.65e+03
Score: 90.50
Percent Similarity: 45.00%
Best Local Similarity: 27.00%
Query Match: 11.57%
DB: 3
Length: 4403765
Matches: 27
Conservative: 18
Mismatch: 22
Indels: 33
Gaps: 4

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-103-840A-2 (1-4403765)


```

; ORGANISM: Enterococcus faecalis
US-09-134-000C-790

Alignment Scores:
Pred. No.: 0.129
Score: 80.00
Length: 978
Percent Similarity: 39.01%
Matches: 37
Best Local Similarity: 26.24%
Conservative: 18
Query Match: 10.23%
Mismatch: 58
Indels: 28
DB: 4
Gaps: 6

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-134-000C-790 (1-978)

QY 9 GlyAsnValThrAsnIle---LeuGluLeuArgAsnLysAspPheArgHisSerHis 27
Db 289 GGCAATGCGATGAAACATACCTGCGCCCATTCGAATTGAAGCATTTGATAATTCAAAT--- 345
QY 28 ProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSerAsp 47
Db 346 -----ATTATGGGAACAAATCCTGTTTCTCGATGGTTGTTTATTGAC 390
QY 48 -----SerAsnSerArgArgValPheLysIleLysSerThrValValLys 63
Db 391 GGACGGCCAGCTAAATGAGTATCGAAATATAAAATATAAACCGTCCCAAGTCTCTGAT 450
QY 64 AspLeuValLysAsnSerGluVal-----AlaGlyThrGlyAspGlnCys 79
Db 451 GATTACGCTTCTATGAGAGAAGTCAATTATCGCAGATATTACAGAGTTTAAAGAAGGC 510
QY 80 LeuProPheAspAspThrArgCysGlyAspGlyGlyLys-----AlaThrGlu 95
Db 511 TTGCTTTTCTGATTAAATTTTAAATCGATGGTGAAGGTCAAGTCGATGTTGCCAAA 570
QY 96 AlaThrLeuThrAsnProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheVal 115
Db 571 GATGTGTTAGCCAATCAACTAGGAGTCGATATTCCTGTAGCTGGATTA----- 618
QY 116 AspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySer 135
Db 619 -----GCCAAGAACGACAAAGCACAAACAGTGAACCTTTTATTGTTGCCA 663
QY 136 Asn 136
Db 664 AAC 666
```

```

RESULT 7
US-09-235-451-3/c
; Sequence 3, Application US/09235451
; GENERAL INFORMATION:
; APPLICANT: Julius, David J.
; APPLICANT: Caterina, Michael J.
; APPLICANT: Brake, Anthony J.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING
; TITLE OF INVENTION: CAPSAICIN RECEPTOR AND CAPSAICIN RECEPTOR-RELATED
; FILE REFERENCE: 9076/084CTP
; CURRENT APPLICATION NUMBER: US/09/235,451
; CURRENT FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/072,151
; PRIOR FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: 08/915,461
; PRIOR FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2736
; TYPE: DNA
; ORGANISM: R. rattus
; FEATURE:
US-09-235-451-3
```

```

Alignment Scores:
Pred. No.: 0.735
Length: 2736
```

```

Score: 79.50
Percent Similarity: 41.38%
Best Local Similarity: 28.97%
Query Match: 10.17%
DB: 4
Matches: 42
Conservative: 18
Mismatch: 79
Indels: 6
Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-235-451-3 (1-2736)

QY 7 ProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLysAspPheArgHisSerHis 26
Db 453 CCCTCTGGAATGTTGACTCCATGGGGGGCGGTTCTCTGCTTCCCTTGTTCACCTCAGCAT 394
QY 27 SerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSer 46
Db 393 TGCCCTCTTCACTCCATCGGAAGTCTCCAGCTGAAAGCTGGGGGCTGAGGCTGAAG 334
QY 47 AspSerAsnSerArgArgValPheLysIleLysSerThrValValLysAspLeuVal 66
Db 333 TCATCCTGAGAGAGGAGGATGCTCCTTACAGCTGAAACCTCTGTAGAGATGATGTTA 274
QY 67 LysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeuProPheAspThrAr 86
Db 273 AGGTTCTCTCTCATACCGCAGGGGCGAGGCTGGTGGGACCCCGAGGCTGAAGCCAAG 214
QY 86 gCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThrVa 106
Db 213 GTGT---AGCGGGCGGCAAA---CTGCAGGCTAGGATCACCGAGCTAGCGGG---GCTCT 163
QY 106 lAspLysPheGlyLeuIleTyrPheVal---AspGlyThrMetIleArgArgIleAspGl 125
Db 162 CCGCAGCGGGGACTGAGAGAGGAAGTGGCGGCGGAGGGGTGCGGCTCAACGCTGACAG 103
QY 125 nAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSerAlaArgProLeuSe 145
Db 102 TAATGGGCTGTGGGAACCTCCGCGGCGGCTCCGAGCTTCTCTGCGTCTGCCACCTC 43
QY 145 r---CysAspSer 148
Db 42 CACCTGTTCGTCT 30
```

```

RESULT 8
US-08-652-877-85
; Sequence 85, Application US/08652877
; Patent No. 6187548
; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjalms, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
```

; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 85:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Placenta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 68..14035
US-08-652-877-85

Alignment Scores:
Pred. No.: 12.6 Length: 14042
Score: 78.50 Matches: 32
Percent Similarity: 41.35% Conservative: 23
Best Local Similarity: 24.06% Mismatches: 54
Query Match: 10.04% Indels: 24
DB: 3 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-652-877-85 (1-14042)

QY 25 SerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPhe 44
Db 2276 TCGGGGAATCCTTCTTTGTGCGGGATTGATTGTGACGCCAGGACAGCACTATCTTT 2335
QY 45 LeuSerAspSerAsnSerArgArgValPheLysIleLys-----SerThr 59
Db 2336 TTTTCAGATATGTCAAAACACATGATTTTAAAGCAAAGATTGATGGCAGGAGAGAA 2395
QY 60 ValValValLysAspLeuValLysAsnSerGluValValAla----- 73
Db 2396 ATTCTCGCAGCTAACAGGGTGGAATGTTGAAAGTTTGGCTTTTGACTGGATTCAAAG 2455
QY 74 -----GlyThrGlyAspGlnCysLeuProPheAspThrArgCysGlyAspGly 90
Db 2456 AATCTCTATTGGACAGACTCTCATTACAAGAGTATCAGTGTATGAGGCTAGCTGATAAA 2515
QY 91 GlyLysAlaThrGluAlaThr---LeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db 2516 ACGAGACGCACGGTAGTTTCAGTATTTAAATAACCCACGGTCGGTGGTAGTTTCATCCTTT 2575
QY 110 ---GlyLeuIleTyrPheValAsp-GlyThrMetIleArgArgIleAspGlnAsnGlyI 128
Db 2576 GCCGGGTATCTATTCTTCACTGATTGGTTCCGTCCTGCTAAATATATGAGACATGGAGT 2635
QY 128 eIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2636 G-----ACGGATCTCACCTCT 2651

RESULT 9

US-08-652-877-89
; Sequence 89, Application US/08652877
; Patent No. 6187548

; GENERAL INFORMATION:
; APPLICANT: Akerstrom, Goran
; APPLICANT: Juhlin, Claes
; APPLICANT: Rask, Lars
; APPLICANT: Crumley, Gregg R.
; APPLICANT: Morse, Clarence C.
; APPLICANT: Murray, Edward M.
; APPLICANT: Hjalms, Goran
; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments
; TITLE OF INVENTION: Thereof and DNA Encoding Same
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Rd., 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.5.1
; SOFTWARE: Word 6.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,877
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/15203
; FILING DATE: 22-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,836
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,314
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Savitzky, Martin
; REGISTRATION NUMBER: 29,699
; REFERENCE/DOCKET NUMBER: A1355E-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-454-3816
; TELEFAX: 610-454-3808
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14044 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Parathyroid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 65..14032
US-08-652-877-89
Alignment Scores:
Pred. No.: 12.6 Length: 14044
Score: 78.50 Matches: 32
Percent Similarity: 41.35% Conservative: 23
Best Local Similarity: 24.06% Mismatches: 54
Query Match: 10.04% Indels: 24
DB: 3 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-652-877-89 (1-14044)

QY 25 SerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPhe 44
Db 2273 TCGGGGAATCCTTCTTTGTGCGGGATTGATTGTGACGCCAGGACAGCACTATCTTT 2332

QY 45 LeuSerAspSerAsnSerArgArgValPheLysIleLys-----SerThr 59
Db 2333 TTTTCAGATATGTCAAAACACATGATTTTAAAGCAAAAGATTGATGGCACAGGAAGAA 2392
QY 60 ValValVallysAspLeuValLysAsnSerGluValValAla----- 73
Db 2393 ATTCTCGCAGCTAACAGGGTGGAATGTTGAAAGTTTGGCTTTTGATGGATTTCAAAG 2452
QY 74 -----GlyThrGlyAspGlnCysLeuPropheAspThrArgCysGlyAspGly 90
Db 2453 AATCTCTATTGGACAGACTCTCATTTACAAGAGTATCAGTGTCTATGAGGCTAGCTGATAAA 2512
QY 91 GlyLysAlaThrGluAlaThr---LeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db 2513 ACGAGACGCACGGTAGTTTCAGTATTTAAATAACCCACGGTCGGTGGTAGTTTCATCCTTTT 2572
QY 110 ---GlyLeuIleTyrPheValAsp-GlyThrMetIleArgArgIleAspGlnAsnGlyI1 128
Db 2573 GCCGGGTATCTATCTTCACTGATTGGTTCGCTCCTGCTAAATTAATGAGAGCATGGAGT 2632
QY 128 eileSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2633 G-----ACGGATCTCACCTCT 2648

RESULT 10

US-08-652-877-87

; Sequence 87, Application US/08652877

; Patent No. 6187548

; GENERAL INFORMATION:

; APPLICANT: Akerstrom, Goran

; APPLICANT: Juhlin, Claes

; APPLICANT: Rask, Lars

; APPLICANT: Crumley, Gregg R.

; APPLICANT: Morse, Clarence C.

; APPLICANT: Murray, Edward M.

; APPLICANT: Hjaln, Goran

; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments

; NUMBER OF INVENTION: Thereof and DNA Encoding Same

; NUMBER OF SEQUENCES: 106

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Rhone-Poulenc Rorer Inc.

; STREET: 500 Arcola Rd., 3C43

; CITY: Collegeville

; STATE: PA

; COUNTRY: USA

; ZIP: 19426-0107

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: Macintosh

; OPERATING SYSTEM: System 7.5.1

; SOFTWARE: Word 6.0 (Patentin)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/652,877

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/15203

; FILING DATE: 22-NOV-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/344,836

; FILING DATE: 23-NOV-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/487,314

; FILING DATE: 07-JUNE-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Savitzky, Martin

; REGISTRATION NUMBER: 29,699

; REFERENCE/DOCKET NUMBER: A1355E-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 610-454-3816

; TELEFAX: 610-454-3808

; INFORMATION FOR SEQ ID NO: 87:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 14080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Kidney
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 105..14072
US-08-652-877-87

Alignment Scores:

Pred. No.: 12.6 Length: 14080
Score: 78.50 Matches: 32
Percent Similarity: 41.35% Conservative: 23
Best Local Similarity: 24.06% Mismatches: 54
Query Match: 10.04% Indels: 24
DB: 3 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-652-877-87 (1-14080)

QY 25 SerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPhe 44

Db 2313 TCGGGGAATCCTTCTTCTTTGTCGGGATGATTTTGACGCCACAGGACACTATCTTT 2372

QY 45 LeuSerAspSerAsnSerArgArgValPheLysIleLys-----SerThr 59

Db 2373 TTTTCAGATATGTCAAAACACATGATTTTAAAGCAAAAGATTGATGGCACAGGAAGAA 2432

QY 60 ValValVallysAspLeuValLysAsnSerGluValValAla----- 73

Db 2433 ATTCTCGCAGCTAACAGGGTGGAATGTTGAAAGTTTGGCTTTTGATGGATTTCAAAG 2492

QY 74 -----GlyThrGlyAspGlnCysLeuPropheAspThrArgCysGlyAspGly 90

Db 2493 AATCTCTATTGGACAGACTCTCATTTACAAGAGTATCAGTGTCTATGAGGCTAGCTGATAAA 2552

QY 91 GlyLysAlaThrGluAlaThr---LeuThrAsnProArgGlyIleThrValAspLysPhe 109

Db 2553 ACGAGACGCACAGTAGTTTCAGTATTTAAATAACCCACGGTCGGTGGTAGTTTCATCCTTTT 2612

QY 110 ---GlyLeuIleTyrPheValAsp-GlyThrMetIleArgArgIleAspGlnAsnGlyI1 128

Db 2613 GCCGGGTATCTATCTTCACTGATTGGTTCGCTCCTGCTAAATTAATGAGAGCATGGAGT 2672

QY 128 eileSerThrLeuLeuGlySerAsnAspLeuThrSer 140

Db 2673 G-----ACGGATCTCACCTCT 2688

RESULT 11

US-08-476-515A-83

; Sequence 83, Application US/08476515A

; Patent No. 6239270

; GENERAL INFORMATION:

; APPLICANT: Akerstrom, Goran

; APPLICANT: Juhlin, Claes

; APPLICANT: Rask, Lars

; APPLICANT: Crumley, Gregg R.

; APPLICANT: Morse, Clarence C.

; APPLICANT: Murray, Edward M.

; APPLICANT: Hjaln, Goran

; TITLE OF INVENTION: Human Calcium Sensor Protein, Fragments

; NUMBER OF INVENTION: Thereof and DNA Encoding Same

; NUMBER OF SEQUENCES: 84

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Martin Savitzky

; STREET: Rhone-Poulenc Rorer Inc., 500 Arcola Rd.;

; STREET: 3C43,

Score: 78.50 Matches: 32
Percent Similarity: 41.35% Conservative: 23
Best Local Similarity: 24.06% Mismatches: 54
Query Match: 10.04% Indels: 24
DB: 3 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-08-652-877-83 (1-14086)

QY 25 SerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPhe 44
Db 2315 TCGGGAAATCCTTCTTCTTGTCTGGGATTTGACGCCAGGACAGCACTATCTTT 2374

QY 45 LeuSerAspSerAsnSerArgValPheLysIleLys-----SerThr 59
Db 2375 TTTTCAGATATGTCAAAACACATGATTTTAAAGCAAAAGATTGATGGCACAGGAAGAA 2434

QY 60 ValValLysAspLeuValLysAsnSerGluValValAla----- 73
Db 2435 ATTCTCGAGCTAACAGGGTGGAAATGTTGAAAGTTTGGCTTTTGATTGGATTTCAAAG 2494

QY 74 -----GlyThrGlyAspGlnCysLeuProPheAspThrArgCysGlyAspGly 90
Db 2495 AATCTCTATTGGACAGACTCTCATTACAAGAGTATCAGTGTCTATGAGGCTAGTGATAAA 2554

QY 91 GlyLysAlaThrGluAlaThr---LeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db 2555 ACGAGACGCACAGTAGTTCAGTATTAAATAACCCCGTGGTGGTAGTTTCATCTTTT 2614

QY 110 ---GlyLeuIleTyrPheValAsp-GlyThrMetIleArgArgIleAspGlnAsnGlyI 128
Db 2615 GCCGGGTATCTTCTTCTACTGATTTGTTCCGTCCTCTGCTAAATATGAGAGCATGGAGT 2674

QY 128 eileSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 2675 G-----ACGATCTCACCTCT 2690

RESULT 13

US-09-453-702B-213/c
; Sequence 213, Application US/09453702B
; Patent No. 6365723

GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000

TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 213:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6943
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 213:
US-09-453-702B-213

Alignment Scores:
Pred. No.: 8.22 Length: 6943
Score: 76.50 Matches: 36
Percent Similarity: 37.34% Conservative: 23
Best Local Similarity: 22.78% Mismatches: 68
Query Match: 9.78% Indels: 31
DB: 4 Gaps: 7

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-453-702B-213 (1-6943)

QY 13 AsnIleLeuGluLeuArgAsnLysAspPheArgHisSerHisSerProAlaHisLysTyr 32
Db 987 AATTATAGCGTTCGTGTAAGTGTGATTTCAAAACATAGCAATGGCGTTATCAGGAAATA 928

QY 33 TyrLeuAlaThrAspProMetSerGlyAlaValPheLeuSerAspSerAsnSerArg 52
Db 927 TATTATACCGCAGACCTGCTCCCGTATGTTTATAGCACCAACCGAAGTGTCTTGTCT 868

QY 53 ValPheLysIleLysSerThrValValLysAspLeuValLysAsnSerGluValVal 72
Db 867 TTTTACTATCTTAACGAATATGTCGATGTGGGAACAAAAATATCT-----GTG 820

QY 73 AlaGlyThrGlyAspGlnCysLeuProPheAsp----- 83
Db 819 CTAAATGCGGGGTATACGGCAGTTCCTTTTGAACATGTTTCCAACCAAGGCACTACAACA 760

QY 84 AspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsnProArgGly 103
Db 759 GATCACACTTGT---CAGGGAACAAACTACAGCGTGGCGTGAGCCTGAAACTGGA 703

QY 104 IleThrValAsp---LysPheGlyLeuIleTyrPheValAspGlyThrMetIleArg 122
Db 702 GCAGATCGGAGATTTCATTTCGTATTAAACGTTCAATAATGGAACGGTAGTAATACCT 643

QY 123 IleAspGlnAsnGlyIleIle-----SerThrLeuLeuGly----- 134
Db 642 ATCACCAGATATTCATGCTGTATGCCAACATATCCAGCACCAACCGCTGTGAGCGG 583

QY 135 -----SerAsnAspLeuThrSerAlaArgProLeuSerCys 146
Db 582 ATTGCAAAAGTTGCAATTTGAGGAGTTTGACCGCA-----CCACAGTCTTGT 535

RESULT 14

US-09-134-001C-2101
; Sequence 2101, Application US/09134001C
; Patent No. 6380370

GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2101
; LENGTH: 2073
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis


```
US-09-134-001C-2101
Alignment Scores:
Pred. No.: 1.52 Length: 2073
Score: 76.00 Matches: 32
Percent Similarity: 33.86% Conservative: 11
Best Local Similarity: 25.20% Mismatches: 32
Query Match: 9.72% Indels: 52
DB: 4 Gaps: 5

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-134-001C-2101 (1-2073)
QY 20 LysAspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPro--- 38
Db 196 AAAGATTTAAGAAAAAACATGCGAAAAAGCAAAAAAATATTTTAGCTAGTGACCTGAC 255
QY 39 -----MetSerGlyAlaValPheLeuSerAspSerAsn 49
Db 256 CGTGAAGGTGAAGCGATTGCTTGGCAATTATCAAAAATTTTAGAATTAGAAGATAGCAAA 315
QY 50 SerArgArgValPheLysIleLysSerThrValValVallysAspLeuVallysAsnSer 69
Db 316 GAAATAGAGTAGTA-----TTTAAATGAAATTACAAAAGATGCTGTTAAAGATAGT 366
QY 70 GluValValAlaGlyThrGlyAspGlnCysLeuPropheAspThrArgCysGlyAsp 89
Db 366 ----- 366
QY 90 GlyGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThrValAspLysPhe 109
Db 367 -----TTTAAGCATCCTCGTGGTATTGAAATGGAT----- 396
QY 110 GlyLeuIleTyrPheValAspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIle 129
Db 397 -----TTAGTTGACGCGCAACAAGCACGTCGT-----ATTTTA 429
QY 130 SerThrLeuLeuGlySerAsn 136
Db 430 GATAGACTCGTGGTTATAAT 450

RESULT 15
US-09-453-702B-244/c
; Sequence 244, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burland, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
```

```
;
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2797
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 244:
US-09-453-702B-244
Alignment Scores:
Pred. No.: 3.34 Length: 2797
Score: 75.00 Matches: 38
Percent Similarity: 36.99% Conservative: 26
Best Local Similarity: 21.97% Mismatches: 59
Query Match: 9.59% Indels: 50
DB: 4 Gaps: 8

US-10-029-020-14_COPY_1250_1400 (1-151) x US-09-453-702B-244 (1-2797)
QY 2 IleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLysAsp 21
Db 1541 ATTGCTAATTCTCCCGGACAATAATCTGACC-----CCAAATGAGAGC 1497
QY 22 PheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGly 41
Db 1496 TGGAAAAGATTTTGTGAGACATGCGAGCGCTTTTACGTTGCTCAGAAGAGTATCCTGGA 1437
QY 42 -----AlaValPheLeuSerAspSerAsnSerArgArgVal 53
Db 1436 GATAAGAGTGAACGTCTTACGGGAAAACTCTCTATCTCTGATGCGGAATTACAATGACC 1377
QY 54 PheLysIleLysSerThrValValVallysAspLeuVallysAsnSerGluValAla 73
Db 1376 TTCAAGATA-----GGTGATGTTGTCTATC 1353
QY 74 GlyThrGlyAspGlnCysLeuProPheAspAspThr-----ArgCysGlyAspGlyGly 91
Db 1352 AATACTATTAGCACTGCTATTCTCGAAGATGCAACGGGTCAACGGTGTATCGAAGGGTTG 1293
QY 92 LysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThrValAspLysPheGlyLeu 111
Db 1292 AATTAGCAGAGATGGATTTAACCGAC-----ATAGACTTGTGAAAAATGGCGCTA 1242
QY 112 -----IleTyrPheValAspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIle 129
Db 1241 AGGAATGTCAATTTT---AATGGCAGCATTTCTTAGAAATGCCAAGTCTCCGGTACGATC 1185
QY 130 -----SerThrLeu 132
Db 1184 TGTGAAGCGGTGGATTTTACCGATTGTGATCTGCGTAATGCGAATTCGAAATGCCTCA 1125
QY 133 LeuGlySerAsnAspLeuThrSerAlaArgProLeuSer 145
Db 1124 TTAGAAAATAATGATTTTTCGTAAAGTTCGCCACTTGACT 1086

Search completed: August 14, 2004, 21:33:24
Job time : 5053.67 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 1770.9 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400
Perfect score: 782
Sequence: 1 YIRRIFFSGVNTNILELRNK.....LLGSNDLTSARPLSCDSVMD 151

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_113135@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	727.5	93.0	6246	29	AY413475	AY413475 Homo sapi
2	712.5	91.1	3190	29	AY413477	AY413477 Mus muscu
3	645.5	82.5	770	14	CB524534	CB524534 UI-M-FY0-
4	618	79.0	623	14	CF538162	CF538162 UI-M-GI0-
5	603	77.1	640	14	CF744767	CF744767 UI-M-GV0-
6	588	75.2	757	14	CD351739	CD351739 UI-M-GI0-
7	564.5	72.2	5069	29	AY405421	AY405421 Pan trogl
8	564.5	72.2	5094	29	AY405420	AY405420 Homo sapi
9	563.5	72.1	5087	29	AY405422	AY405422 Mus muscu
10	546.5	69.9	778	13	BU705631	BU705631 UI-M-FO0-
11	540.5	69.1	820	13	BU226585	BU226585 603800545
12	499	63.8	801	14	CD803099	CD803099 UI-M-GV0-
13	484	61.9	692	14	CF530821	CF530821 UI-M-GH0-
14	459	58.7	724	14	CD352566	CD352566 UI-M-GI0-
15	435	55.6	789	14	CD803315	CD803315 UI-M-GV0-
16	408	52.2	975	14	CD326096	CD326096 AGENCOURT
17	401	51.3	428	28	AZ040891	AZ040891 RPCI-23-2
18	398	50.9	443	10	BF950887	BF950887 RC3-NN118
19	339	43.4	351	13	C82439	C82439 C82439 rabb
20	339	43.4	351	13	C83295	C83295 C83295 rabb
21	297	38.0	784	29	BX144542	BX144542 Danio rer
22	285	36.4	1039	12	BM804083	BM804083 AGENCOURT
23	282	36.1	452	13	BY263427	BY263427 BY263427
24	281	35.9	319	28	AQ552763	AQ552763 RPCI-11-4
25	264	33.8	854	14	CF745232	CF745232 UI-M-GV0-
26	256	32.7	843	13	BU262834	BU262834 603818972
27	242	30.9	5970	29	AY413476	AY413476 Pan trogl
28	223	28.5	733	29	CNS05EJ1	AL333766 Tetraodon
29	221.5	28.3	1005	29	CNS04W60	AL309969 Tetraodon
30	211	27.0	1071	29	CNS058S2	AL326348 Tetraodon
31	210	26.9	640	28	AQ896853	AQ896853 HS 3144 A
32	202	25.8	497	28	BZ902852	BZ902852 CH240_23D
33	202	25.8	601	13	BX504880	BX504880 DKFZp686G
34	184	23.5	662	29	CE057449	CE057449 tigr-gss-
35	183	23.4	723	29	AG030576	AG030576 Pan trogl
36	167	21.4	645	28	AZ413372	AZ413372 IM0197C13
37	166	21.2	1461	28	CC289028	CC289028 CH261-170
38	163	20.8	794	14	CB518529	CB518529 UI-M-GH0-
39	141	18.0	360	9	AV191089	AV191089 AV191089
40	139	17.8	601	13	BX509644	BX509644 DKFZp686N
41	138	17.6	602	9	AL710647	AL710647 DKFZp686M
42	111.5	14.3	595	28	BH329569	BH329569 CH230-106
43	109	13.9	532	12	BI418126	BI418126 LjNEST23a
44	105.5	13.5	610	13	BU027852	BU027852 QHG8D16.Y
45	100	12.8	361	14	FL4403	FL4403 ATTS5325 AC

ALIGNMENTS

RESULT 1
AY413475
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE

AY413475
Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY413475
AY413475.1 GI:39769437
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 6246)

6246 bp
DNA
linear
GSS 12-DEC-2003

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 6246)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..6246 Length: 6246

1..6246 Matches: 144

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

gene <1..>6246

/locus_tag="HCM4903"

ORIGIN

Alignment Scores:

Pred. No.: 1.63e-79 Length: 6246

Score: 727.50 Matches: 144

Percent Similarity: 95.36% Conservative: 0

Best Local Similarity: 95.36% Mismatches: 0

Query Match: 93.03% Indels: 7

DB: 29 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x AY413475 (1-6246)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 1699 TACATTAGAGGATCTTCCCTCTGGAAATGTCACCAACATCCTAGAGCTG----- 1749

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40

Db 1750 -----AGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGT 1797

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 1798 GGGGCCGTCTTCCTTCTGACAGCAACAGCCGGGGTCTTTAAAAATCAAGTCCACTGTG 1857

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80

Db 1858 GTGGTGAAGGACCTGTCAAGAACTCTGAGGTGTTGCGGGACAGGTGACCAAGTGCCTC 1917

QY 81 ProPheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100

Db 1918 CCCTTTGATGACACTCGCTGCGGGATGTTGGGAAGGCCACAGAAGCCACACTCACCAAT 1977

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120

Db 1978 CCCAGGGGCATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGATC 2037

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140

Db 2038 AGACGCATCGATCAGAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATCA 2097

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151

Db 2098 GCCCGGCCACTCAGCTGTGATTCTGTCAATGGAT 2130

RESULT 2

AY413477 3190 bp DNA linear GSS 12-DEC-2003

LOCUS AY413477

DEFINITION Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

AY413477

VERSION AY413477.1 GI:39769439

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3190)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 3190)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..3190

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

gene <1..>3190

/locus_tag="HCM4903"

ORIGIN

Alignment Scores:

Pred. No.: 4.65e-78 Length: 3190

Score: 712.50 Matches: 138

Percent Similarity: 94.70% Conservative: 5

Best Local Similarity: 91.39% Mismatches: 1

Query Match: 91.11% Indels: 7

DB: 29 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x AY413477 (1-3190)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 1699 TACATCCGCAGATCTTCCCTCTGGAATGTCACCAACATCTCTGGAGATG----- 1749

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40

Db 1750 -----AGTCACAGCCCGACACACAAATACTACCTGGCTACAGACCCCATGAGT 1797

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 1798 GGGGCCGTCTTCTGTGACACCAACAGCCGGGGTCTTCAAGGTCAAGTCCACCACA 1857

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80

Db 1858 GTGGTGAAGGACCTGGTGAAGAACTCCGAGGTGGTAGCAGGACTGGTGACCGCTC 1917

QY 81 PropheAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100

Db 1918 CCCTTTGATGATACCCGCTCGGAGATGTTGGGAAAGCCACAGAAGCCACGCTCACTAAC 1977

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120

Db 1978 CCCAGGGGAATTACAGTGGACAAGTTTGGGCTCATTTATTTCTGGACGGCACCATGATC 2037

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140

Db 2038 AGACGTGTTGATCAAAATGGAATCATCTCCACTTTGCTGGCTCCAATGACCTCACCTCG 2097

```
QY      141 AlaArgProLeuSerCysAspSerValMetAsp 151
      |||
Db      2098 GCCAGGCCCTCAGCTGTGACTCCGTCATGGAG 2130

RESULT 3
CB524534
LOCUS
DEFINITION
      770 bp      mRNA      linear      EST 09-JUL-2003
      UI-M-FY0-cey-b-20-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone
      IMAGE:6843669 5', mRNA sequence.
CB524534
VERSION
KEYWORDS
SOURCE
      Mus musculus (house mouse)
ORGANISM
      Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
      1 (bases 1 to 770)
REFERENCE
      NIH-MGC http://mgc.nci.nih.gov/.
      National Institutes of Health, Mammalian Gene Collection (MGC)
      Unpublished (1999)
      Contact: Robert Strausberg, Ph.D.
      Email: cgapbs-r@mail.nih.gov
      Tissue Procurement: Dr. Jim Lin, University of Iowa
      cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
      DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
      DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
      Clone Distribution: Distribution information can be found at
      http://genome.uiowa.edu/distribution/mousefl.html
      This clone was contributed by the Brain Molecular Anatomy Project
      (BMAP)
Seq primer: pYX-5.
      Location/Qualifiers
      1. .770
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6"
      /db_xref="taxon:10090"
      /clone="IMAGE:6843669"
      /tissue_type="whole brain"
      /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
      /lab_host="DH10B (T1 phage resistant)"
      /clone_lib="NIH_BMAP_FY0"
      /note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
      Site 2: Not I; The library was constructed according
      Bonaldo, Lennon and Soares, Genome Research, 6:791-806.
      1996. Denatured RNA was size fractionated on a 1% agarose
      gel. First strand cDNA synthesis was primed with oligo-dT
      primer containing a Not I site. Double strand cDNA was
      size selected according to mRNA size fraction, ligated
      with EcoR I adaptor, digested with NotI and then cloned
      directionally into pYX-Asc vector. The library tag
      sequence located between the Not I site and the polyA tail
      is AGCGAGACAG. This library was created for the University
      Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
      Developing Mouse Nervous System', supported by National
      Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
      program coordinator."
ORIGIN
Alignment Scores:
Pred. No.:      1.55e-70      Length:      770
Score:      645.50      Matches:      128
Percent Similarity:      93.66%      Conservative:      5
Best Local Similarity:      90.14%      Mismatches:      1
Query Match:      82.54%      Indels:      8
DB:      14      Gaps:      1

US-10-029-020-14_COPY_1250_1400 (1-151) x CB524534 (1-770)

QY      11 ValThrAsnIleLeu-GluLeuArgAsnLysAspPheArgHisSerHisSerProAlaHi 30
      |||
Db      1 GTCACCAACATCCTNGGAGATG-----AGTCACAGCCCGACACA 39
```

```
QY      30 sLysTyTrLeuAlaThrAspProMetSerGlyAlaValPheLeuSerAspSerAsnSe 50
      |||
Db      40 CAAATACTACCTGGCTACAGACCCCATGAGTGGGCGGCTTCTTCTGTCTGACACCAACAG 99

QY      50 rArgArgValPheLysIleLysSerThrValValValLysAspLeuValLysAsnSerGl 70
      |||
Db      100 CCGGCGGGTCTTCAAGGTCAAGTCCACCACAGTGGTGAAGGACCTGGTGAAGAACTCCGA 159

QY      70 uValValAlaGlyThrGlyAspGlnCysLeuProPheAspAspThrArgCysGlyAspGl 90
      |||
Db      160 GGTGGTAGCAGGGACTGGTGACCAAGTGCCTCCCTTTGATGATACCCCGTCGCGAGATGG 219

QY      90 yGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThrValAspLysPheGl 110
      |||
Db      220 TGGGAAAGCCACAGAAAGCCACGCTCACTAACCCAGGGGAATTACAGTGGACAAGTTTGG 279

QY      110 yLeuIleTyPheValAspGlyThrMetIleArgArgIleAspGlnAsnGlyIleIleSe 130
      |||
Db      280 GCTCATTATTTCGTGGACGGCACCATGATCAGACGCTGTTGATCAAAATGAATCATCTC 339

QY      130 rThrLeuLeuGlySerAsnAspLeuThrSerAlaArgProLeuSerCysAspSerValMe 150
      |||
Db      340 CACTTTGCTGGCTCCAATGACCTCACTCGGCCAGGCCCTCAGCTGTGACTCCGTCAT 399

QY      150 tAsp 151
      |:::
Db      400 GGAG 403

RESULT 4
CF538162
LOCUS
DEFINITION
      623 bp      mRNA      linear      EST 12-SEP-2003
      UI-M-GIO-cho-h-21-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
      IMAGE:30536156 5', mRNA sequence.
CF538162
VERSION
KEYWORDS
SOURCE
      Mus musculus (house mouse)
ORGANISM
      Mus musculus
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
      1 (bases 1 to 623)
REFERENCE
      NIH-MGC http://mgc.nci.nih.gov/.
      National Institutes of Health, Mammalian Gene Collection (MGC)
      Unpublished (1999)
      Contact: Robert Strausberg, Ph.D.
      Email: cgapbs-r@mail.nih.gov
      Tissue Procurement: Dr. Jim Lin, University of Iowa
      cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
      DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
      Clone Distribution: Distribution information can be found at
      http://genome.uiowa.edu/distribution/mousefl.html
      This clone was contributed by the Brain Molecular Anatomy Project
      (BMAP)
Seq primer: pYX-5.
      Location/Qualifiers
      1. .623
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6"
      /db_xref="taxon:10090"
      /clone="IMAGE:30536156"
      /tissue_type="whole brain"
      /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
      /lab_host="DH10B (T1 phage resistant)"
      /clone_lib="NIH_BMAP_G10"
      /note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
      Site 2: Not I; The library was constructed according
      Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
      1996. Denatured RNA was size fractionated on a 1% agarose
      gel. First strand cDNA synthesis was primed with oligo-dt
      primer containing a Not I site. Double strand cDNA was
```


Query Match:	72.19%	Indels:	7		
	29	Gaps:	1		
-10-029-020-14_COPY_1250_1400 (1-151) x AY405421 (1-5069)					
1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20			
872	TATGTGGCGGGATATTCCCTTCTGGAAATGTAACAAGTCTTAGAACTA-----	922			
21	AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer	40			
923	-----AGCAGCAACCCAGCTCATAGATACTACCTTGCACGGATCCAGTCACA	970			
41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60			
971	GGAGATTGTACGTTCTGTACACAAACACCCGCAGAAATTTATCGCCCAAGTCAC	1030			
61	ValVallLysAspLeuValLysAsnSerGluValAlaGlyThrGlyAspGlnCysLeu	80			
1031	GGGGCAAAAGACTTGACTTAAATGCAAGATTGTGCGGGGACAGGGAGCAGTGCCT	1090			
81	ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn	100			
1091	CCGTTTGACGAGCGAGATGTGGGATGGAGGAAGGCCGTGGAAAGCCACACTCAT	1150			
101	ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle	120			
1151	CCCAAGGAATGGCAGTTGATAAGATGGATTAACTACTTTTGTGATGGAACCATG	1210			
121	ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer	140			
1211	AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTGAC	1270			
141	AlaArgProLeuSerCysAspSerValMet	150			
1271	GCCAGACCTTTAACTTGTGACACCAGCATG	1300			
RESULT 8	AY405420	5094 bp	DNA linear GSS 12-DEC-2003		
LOCUS	Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,				
DEFINITION	genomic survey sequence.				
ACCESSION	AY405420				
VERSION	AY405420.1	GI:39761394			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS	1 (bases 1 to 5094) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 5094) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
AUTHORS	Direct Submission				
TITLE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
SOURCE	1..5094 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"				

Db	620	AGAAAGGTTGATCAAAATGGAATCATATCAACTCTCTGNGCTCCAACGACCTCACGTCA	679
QY	141	AlaArgProLeuSerCysAspSerValMet	150
Db	680	GCTCGACCTTTAAGCTGTGATAGTAGCATG	709
RESULT 11			
BU226585			
LOCUS	BU226585	820 bp	mRNA
DEFINITION	603800545F1 CSEQCHN23 Gallus gallus cdna clone	ChEST771i12 5', mRNA	
sequence.			
ACCESSION	BU226585		
VERSION	BU226585.1	GI:25463660	
KEYWORDS	EST.		
SOURCE	Gallus gallus (chicken)		
ORGANISM	Gallus gallus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
AUTHORS	1 (bases 1 to 820)		
TITLE	Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.		
JOURNAL	A Comprehensive Collection of Chicken cDNAs		
MEDLINE	Curr. Biol. 12 (22), 1965-1969 (2002)		
PUBMED	22335534		
COMMENT	12445392		
Contact:	Simon Hubbard		
Department of Biomolecular Sciences			
University of Manchester Institute of Science and Technology			
(UMIST)			
PO Box 88, Manchester, M60 1QD, UK			
Tel: 01612008930			
Fax: 01612360409			
Email: Simon.Hubbard@umist.ac.uk.			
Location/Qualifiers			
1. .820			
/organism="Gallus gallus"			
/mol_type="mRNA"			
/strain="White Leghorn, Hisex"			
/db_xref="taxon:9031"			
/clone="ChEST771i12"			
/dev_stage="22"			
/lab_host="DH10B"			
/clone_lib="CSEQCHN23"			
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1: EcoRI; Site_2: NotI; This normalized library was constructed from 1 million independent clones. CDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."			
ORIGIN			
Alignment Scores:			
red. No.:	2.93e-57	Length:	820
core:	540.50	Matches:	102
ercent Similarity:	86.00%	Conservative:	27
est Local Similarity:	68.00%	Mismatches:	14
uery Match:	69.12%	Indels:	8
B:	13	Gaps:	1
US-10-029-020-14_COPY_1250_1400 (1-151) x BU226585 (1-820)			
QY	1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20

Db	120	TATGTCGACGGATATTTCCATCTCTGGAAATGTAACCTAGTGTCTCGAGCTA	170
QY	21	AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer	40
Db	171	-----AGCAACACACCCAGCTCACAGATACTACCTCGCTACTGACCCCGTGACA	218
QY	41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60
Db	219	GGAGA-TTGTATGTTTCTGACACAAACACGCGCAGGATCTACCGGCCAGGTCACTT	277
QY	61	ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu	80
Db	278	GGTGCCAAAGACCTGACGAAGACGACGAGGTGGTGGAGGGACGGGGAGCAGTGCCTG	337
QY	81	ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn	100
Db	338	CCCTTTGATGAAGCACCGGTGTGGTGATGGGGCAAGCGGTGCGAGGCAAACTCTCATG	397
QY	101	ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle	120
Db	398	CCTAAAGGAATAGCAATTGACAAAGATGGGTTAATCTACTTTTGTGATGGGACCATGATC	457
QY	121	ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer	140
Db	458	AGAAAAGTGGATCAAAATGGAATATATATCCACGTTGCTTGGCTCCAATGATCTGACCTCA	517
QY	141	AlaArgProLeuSerCysAspSerValMet	150
Db	518	GCACGACCTCTGACCTGTGATACCAGCATG	547
RESULT 12			
CD803099			
LOCUS	CD803099	801 bp	mrna
DEFINITION	UI-M-GV0-chn-c-06-0-UI.r1 NIH_BMAP_GV0 Mus musculus CDNA clone	linear	EST 15-JUL-2003
	IMAGE:30543317 5', mRNA sequence.		
ACCESSION	CD803099		
VERSION	CD803099.1	GI:32461925	
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 801)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
Contact:	Robert Strausberg, Ph.D.		
Email:	cgapbs-r@mail.nih.gov		
Tissue Procurement:	Dr. Jim Lin, University of Iowa		
CDNA Library preparation:	Dr. M. Bento Soares, University of Iowa		
CDNA Library Arrayed by:	Dr. M. Bento Soares, University of Iowa		
DNA Sequencing by:	Dr. M. Bento Soares, University of Iowa		
Clone Distribution:	Distribution information can be found at		
http://genome.uiowa.edu/distribution/mousefl.html			
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)			
Seq primer:	pYX-5.		
Location/Qualifiers			
1. .801			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57BL/6"			
/db_xref="taxon:10090"			
/clone="IMAGE:30543317"			
/tissue_type="whole brain"			
/dev_stage="1,5, and 15 days newborn"			
/lab_host="DH10B (T1 phage resistant)"			
/clone_lib="NIH_BMAP_GV0"			
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according			
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,			
1996. Denatured RNA was size fractionated on a 1% agarose			
gel. First strand cDNA synthesis was primed with oligo-dT			

primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag is CGAAGTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 4.82e-52 Length: 801
Score: 499.00 Matches: 90
Percent Similarity: 87.30% Conservative: 20
Best Local Similarity: 71.43% Mismatches: 16
Query Match: 63.81% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x CD803099 (1-801)

QY 26 HisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSerGlyAlaValPheLeu 45
Db 2 AACAGCCAGGACACAAGTATTACTTGGCTGTGGACCCCGTGACCGGCTCACTCTAGTC 61
QY 46 SerAspSerAsnSerArgArgValPheLysIleLysSerThrValValLysAspLeu 65
Db 62 TCTGACACCAACAGTCGCCGAATCTACCGAGTCAAGTCTCTGAGCGGAGCCAAAGACCTG 121
QY 66 ValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeuProPheAspThr 85
Db 122 GCTGGAAATTCGGAAGTTGTGGCAGGACTGGCGAACAATGTCTACCCCTTTGATGAAGCC 181
QY 86 ArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsnProArgGlyIleThr 105
Db 182 CGCTGTGGGATGGAGGAGGCTGTGGACGCCACCCTCATGAGCCCCAGAGGTATTGCA 241
QY 106 ValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIleArgArgIleAspGln 125
Db 242 GTAGACAAGAATGGGCTTATGTACTTTGTTGATGCCACCATGATCCGGAAGGTGGACCAA 301
QY 126 AsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSerAlaArgProLeuSer 145
Db 302 AACGGAATCATCTCCACCCCTGTGGCTCCAATGACCTCAGACTGTCGACCACTGAGC 361
QY 146 CysAspSerValMetAsp 151
Db 362 TGTGACTCGAGCATGGAC 379

RESULT 13
CF530821

LOCUS UI-M-GH0-cgu-g-06-0-UI.r1 NIH_BMAP_GH0 Mus musculus cDNA clone
DEFINITION IMAGE:30356597 5', mRNA sequence.

ACCESSION CF530821
VERSION CF530821.1 GI:34582785
KEYWORDS EST.

SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 692)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at

FEATURES

source

Seq primer: pYX-5.

Location/Qualifiers
1..692
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30356597"
/tissue_type="Whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GH0"
/note="Organ: Brain; Vector: pYX- Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 3.01e-50 Length: 692
Score: 484.00 Matches: 91
Percent Similarity: 85.71% Conservative: 23
Best Local Similarity: 68.42% Mismatches: 17
Query Match: 61.89% Indels: 2
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x CF530821 (1-692)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 297 TACGTCCGGCGGATATTCCCGTCTGGGAATGTGACAAGTGTTTAGAACTAAGAAATAA 356
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrIleuAlaThrAspPrometSer 40
Db 357 GATTTTAGACATAGTAGCAACCCAGCTCACAGATACCTGGCTACGGACCCAGTCACC 416
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 417 GGAGATTGTACGTCCTGATACTAACACCCGAGAAATCTATCGGCCGAAATCACTCAG 476
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 477 GGAGCCAAAGACCTGACTAAAACGCTGAAGTGGTGGCAGGGACCGGGGAACAGTGCCT 536
QY 81 -PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAs 100
Db 537 CCCTTTTGACGAGCCAGGTGTGGGATGGAGGCAAGGCTGTGGAAGCAACGCTCATGAG 596
QY 100 nProArgGlyIleThrValAspLysPheGlyLeuIleTyr-PheValAspGlyThrMetI 120
Db 597 TCNCANAGGAATGGCAATCGATAAGAACCGGACTGATCTACTTTTGTGATGGAACCATGA 656
QY 120 leArgArgIleAspGlnAsnGlyIleIleSerThr 131
Db 657 TCAGAAAGGTTGATCANNATGGAATCATATCAACT 691

RESULT 14

CD352566

LOCUS

CD352566 724 bp mRNA linear EST 09-JUL-2003

[illegible]

Query Match:	55.63%	Indels:	2
DB:	14	Gaps:	0
US-10-029-020-14_COPY_1250_1400 (1-151) x CD803315 (1-789)			
QY	1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20
Db	398	TACGTCCGGCGGATATTCCCGTCTGGGAATGTGACAAAGTGTTTAGAACTAAGAAATAAA	457
QY	21	AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer	40
Db	458	GATTTTAGACATAGTAGCAACCCAGCTCACAGATACTACCTGGCTACGGACCCAGTCACC	517
QY	41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60
Db	518	GGAGATTTGTACGTCTCTGATACTAACACCCGACAGATCTATCGGCCGAAATCACTCACG	577
QY	61	ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu	80
Db	578	GGAGCCAAAGACCTGACTAAACCGCTGAAGTGGTGGCAGGGACCGGGGAACAGTGCCTT	637
QY	81	ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn	100
Db	638	CCCTTTGACGAGGCCAGGTGTGGGGATGGAGGCAAGGCTGTGGAAGC-ACGCTCATGAGT	696
QY	101	ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle	120
Db	697	CNCAAGGAATGGCAATCGATAAGAACGGACTGATCTACTTTGTGTATGGACC-ATGATC	755
QY	121	ArgArgIle	123
Db	756	AGAAGGTTG	764

Search completed: August 14, 2004, 18:03:23
Job time : 1784.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 256.137 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1250_1400
Perfect score: 782
Sequence: 1 YIRRIFFSGNVNILELRN.....LLGSNDLTSARPLSCDSVMD 151

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues
Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN 1 1 1868 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	782	100.0	8354	6	ABS52100 Human TEN
2	759.5	97.1	8438	6	ABN85378 Human NOV
3	727.5	93.0	8645	6	ABS78652 Human CDN
4	619	79.2	8473	6	ABQ82345 Human NOV
5	619	79.2	8487	6	ABQ82346 Human NOV
6	619	79.2	8645	6	ABQ82344 Human NOV
7	619	79.2	8675	6	ABQ82343 Human NOV
8	613	78.4	9058	7	ACC72051 BCU0205A

9	613	78.4	9695	7	ACC72052
10	613	78.4	9729	5	AAS14089
11	613	78.4	9729	9	ADB32028
12	613	78.4	9826	5	AAS14085
13	613	78.4	9826	9	ADB32023
14	571.5	73.1	4245	7	AAL60066
15	534.5	68.4	12879	6	ABK92230
16	534.5	68.4	13202	4	AAK51828
17	281.5	36.0	10242	4	ABL29075
18	281.5	36.0	17131	4	ABL29074
19	253	32.4	1764	4	ABL04855
20	253	32.4	4925	4	ABL04854
21	101	12.9	1482	6	ABZ14144
22	91	11.6	247	3	AAK18884
23	90.5	11.6	110000	4	AAI99682_10
24	90.5	11.6	110000	4	AAI99683_10
25	87.5	11.2	1842	3	AAC39026
26	85	10.9	2550	9	ADB61918
27	85	10.9	2700	6	ABN59835
28	85	10.9	4308	6	ABS67814
29	84	10.7	6437	4	AAK82889
30	83	10.6	110000	4	AAI99682_39
31	83	10.6	110000	4	AAI99683_39
32	82	10.5	4377	4	AAS55933
33	81	10.4	4389	7	ABX05900
34	81	10.4	9607	2	AAV52213
35	81	10.4	15438	7	ABT42143
36	81	10.4	110000	7	ABS56454_01
37	79.5	10.2	2736	2	AAK87478
38	79.5	10.2	2736	2	AAI9730
39	79	10.1	1610	2	AAI13498
40	79	10.1	1610	6	ABS99293
41	78.5	10.0	14042	2	AAT28774
42	78.5	10.0	14042	2	AAV05995
43	78.5	10.0	14044	2	AAT28776
44	78.5	10.0	14044	2	AAV05997
45	78.5	10.0	14080	2	AAT28775

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
XX
AC ABS52100;
XX

DT 05-NOV-2002 (first entry)

XX Human TEN-M4-like gene.

XX Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophilin;
KW type 1a membrane-sushi domain containing; SNP; gene; ds;
single nucleotide polymorphism.

OS Homo sapiens.

XX Key Location/Qualifiers
FH variation replace(117,G)
FT /*tag= a

FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT /replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
PD 25-JUL-2002.
XX
PF 19-DEC-2001; 2001WO-US050331.
XX
PR 19-DEC-2000; 2000US-0256704P.
PR 20-DEC-2000; 2000US-0257314P.
PR 02-MAY-2001; 2001US-0288153P.
PR 29-MAY-2001; 2001US-0294075P.
PR 24-JUL-2001; 2001US-0307506P.
PR 10-AUG-2001; 2001US-0311590P.
PR 10-AUG-2001; 2001US-0311613P.
PR 29-AUG-2001; 2001US-0315617P.
PR 14-SEP-2001; 2001US-0322358P.
XX
PA (CURA-) CURAGEN CORP.
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR P-PSDB; ABG70388.
XX
PT Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.
XX
PS Claim 8; Page 50-52; 318pp; English.
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders. The
CC present nucleic acid sequence represents a NOVX gene. This sequence
CC encodes a NOVX protein of the invention
XX
SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.99e-95 Length: 8354
Score: 782.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ABS52100 (1-8354)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3782 TACATTAGAAGGATCTTCCCTCTGGAAATGTACCAACATCCTAGAGCTGAGGAATAA 3841

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3842 GATTTCAGACATAGTCACAGTCCAGCACACAATACTACCTGGCCACAGACCCCATGAGT 3901
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3902 GGGGCCGTCTTCCCTTCTGACAGCAACAGCCGGGGGTCTTTAAATCAAGTCCACTGTG 3961
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3962 GTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCCGGGACAGGTGACCATGCTC 4021
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4022 CCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCCACACTCACAAT 4081
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4082 CCCAGGGGTATTACAGTGGACAAAGTTGGGCTGATCTACTTCTGGATGGCACCATGATC 4141
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4142 AGACGCATCGATCAGAATGGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATCA 4201
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4202 GCCCGGCCACTCAGCTGTGATTTCTGTCTATGGAT 4234
RESULT 2
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
XX
AC ABN85378;
XX
DT 21-OCT-2002 (first entry)
XX
DE Human NOV1, TEN-M4 like protein, coding sequence.
XX
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 4..8395
FT /*tag= a
FT /trans_except= (pos: 1138..1147,aa:Met)
FT /product= "NOV1 protein"
XX
PN WO200255704-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-US0000554.
XX
PR 09-JAN-2001; 2001US-0260417P.
PR 10-JAN-2001; 2001US-0260831P.
PR 28-FEB-2001; 2001US-0272338P.
PR 09-MAR-2001; 2001US-0274876P.
PR 18-APR-2001; 2001US-0284704P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;

Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
 Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
 Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
 Gunther E, Stone DJ;
 WPI; 2002-590674/63.
 P-PSDB; ABB98401.
 NOVX polypeptides and encoding polynucleotides, useful for preventing or
 treating NOVX-associated disorders e.g. cancer, inflammation, or
 Alzheimer's disease, and in chromosome mapping, tissue typing or
 pharmacogenomics.
 Claim 9; Page 8-9; 358pp; English.
 The present sequence is a coding sequence for a NOV protein. The NOV
 proteins and coding sequences are useful for treating or preventing NOV-
 associated disorders or in the manufacture of a medicament for treating
 the disorders, such as cancer, heart disease, inflammation, autoimmune
 disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 IgA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
 (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 and other wasting disorders associated with chronic diseases. NOV1 is a
 TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
 Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

ABS78652;

16-DEC-2002 (first entry)

Human cDNA encoding CGDD10, INCYTE 7488573CB1.

Human; ss; gene; cell growth; differentiation; death; CGDD; cancer; cell proliferative disorder; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis; polycythaemia vera; primary thrombocytopaenia; developmental disorder; renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS; neurological disorder; Alzheimer disease; Parkinson's disease; asthma; reproductive disorder; infertility; autoimmune disorder; gout; allergy; inflammatory disorder; acquired immunodeficiency syndrome; uveitis; autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection; diabetes mellitus; glomerulonephritis; irritable bowel syndrome; multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis; rheumatoid arthritis.

Homo sapiens.

WO200272830-A2.

19-SEP-2002.

08-FEB-2002; 2002WO-US003715.

09-FEB-2001; 2001US-0268111P.

23-FEB-2001; 2001US-0271175P.

08-MAR-2001; 2001US-0274503P.

09-MAR-2001; 2001US-0274552P.

(INCY-) INCYTE GENOMICS INC.

Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR, Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N, Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;

WPI; 2002-723356/78.

P-PSDB; ABG97359.

New human proteins associated with cell growth, differentiation and death, useful for diagnosing, treating or preventing autoimmune or inflammatory disorders (e.g. AIDS, allergy or anaemia), cancer, atherosclerosis or hepatitis.

Claim 5; Page 175-178; 181pp; English.

The invention relates to an isolated polypeptide comprising CGDD1-12 (cell growth, differentiation and death), a naturally occurring amino acid sequence at least 90% identical to CGDD, a biologically active fragment or an immunogenic fragment. Also included are the polynucleotides encoding CGDD1-12, a recombinant polynucleotide comprising a promoter sequence operably linked to the CGDD polynucleotides, a cell transformed with the recombinant polynucleotide, a transgenic organism comprising the recombinant polynucleotide, an anti-CGDD antibody, screening for compounds which bind to/modulate or are ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a CGDD polynucleotide microarray. The polypeptides, polynucleotides, agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of CGDD, particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary thrombocytopaenia or cancer), developmental disorders (e.g. renal tubular acidosis, anaemia or mental retardation), neurological disorders (e.g. Alzheimer disease, Parkinson's disease or epilepsy), reproductive disorders (e.g. infertility or a disruption in the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS, (acquired immunodeficiency syndrome) allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus, glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis, uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic

CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2.21e-87 Length: 8645
Score: 727.50 Matches: 144
Percent Similarity: 95.36% Conservative: 0
Best Local Similarity: 95.36% Mismatches: 0
Query Match: 93.03% Indels: 7
DB: 6 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x ABS78652 (1-8645)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3852 TACATTAGAAGGATCTTCCCTCTGGAAATGTCAACCAATCCTAGAGCTG----- 3902

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3903 -----AGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAGT 3950

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3951 GGGGCGTCTTCTTCTGACAGCAACAGCGGGGCTCTTTAAATAATCAAGTCCACTGTG 4010

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4011 GTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGTTGCGGGGACAGGTGACCACTGCCTC 4070

QY 81 PropheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4071 CCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAGCCACACTCACCAAT 4130

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4131 CCCAGGGGCATTACAGTGGACAAAGTTTGGCTGATCTACTTCGTGGATGGCACCATGATC 4190

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4191 AGACGCATCGATCAGAATGGGATCATCTCACCCCTGCTCGGCTCTAATGATCTCACATCA 4250

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4251 GCCCGGCCACTCAGCTGTGATTCTGTGCATGGAT 4283

RESULT 4
ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX
AC ABQ82345;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15c encoding cDNA SEQ ID NO:39.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; allergic reaction; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.
OS
XX

FH Key Location/Qualifiers
FT CDS 258..8144
FT /*tag= a
FT /product= "NOV15c"
XX
PN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53588.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 119-121; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or goitre), and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15c, which is
CC located on chromosome 4
XX
SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 1.27e-72 Length: 8473

Score:	619.00	Matches:	111
Percent Similarity:	90.67%	Conservative:	25
Best Local Similarity:	74.00%	Mismatches:	14
Query Match:	79.16%	Indels:	0
DB:	6	Gaps:	0
US-10-029-020-14_COPY_1250_1400 (1-151) x ABQ82345 (1-8473)			
QY	1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20
Db	3600	TACGTGCGGGATATCCCTTCTGGAAATGTAAACAAGTGTCTTAGAACTAAGAAATAAA	3659
QY	21	AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer	40
Db	3660	GATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG	3719
QY	41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60
Db	3720	GGAGATCTGTACGTTTCTGACACAAACCCGAGAAATTATCGCCCAAGTCACCTACG	3779
QY	61	ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu	80
Db	3780	GGGGCAAAGACTTGACTAAATAATGCAGAACTCGTCGCAGGGACAGGGAGCAATGCCTT	3839
QY	81	PropheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn	100
Db	3840	CCGTTTGCAGGCGAGATGTGGGGATGGAGGAAGCCGTGGAAGCCACACTCATGAGT	3899
QY	101	ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetile	120
Db	3900	CCCAAAGGAATGGCAGTTGATAAGATGGAATGAATTAATCTACTTTGTTGATGGAACCATGATT	3959
QY	121	ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer	140
Db	3960	AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTGTGACTTCA	4019
QY	141	AlaArgProLeuSerCysAspSerValMet	150
Db	4020	GCCAGACCTTTAACTTTGTGACACCCAGCATG	4049
RESULT 5			
ABQ82346			
ID	ABQ82346	standard; cDNA; 8487 BP.	
AC	ABQ82346;		
XX			
XX			
DT	17-DEC-2002	(first entry)	
DE	Human NOV15d	encoding cDNA SEQ ID NO:41.	
XX			
KW	Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;		
KW	cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;		
KW	antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;		
KW	antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;		
KW	cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;		
KW	stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;		
KW	Pick's disease; vesicular transport disease; cystic fibrosis; goitre;		
KW	diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;		
KW	ulcerative colitis; gastric disorder; duodenal disorder; infection;		
KW	autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;		
KW	rheumatoid arthritis; gene; chromosome 4; ss.		
OS	Homo sapiens.		
XX			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	299..8140	
FT		/*tag= a	
FT		/product= "NOV15d"	
XX			
XX	WO2002622999-A2.		
XX			
XX	15-AUG-2002.		
PD			
XX			

PF	31-DEC-2001; 2001WO-US049976.
XX	
PR	29-DEC-2000; 2000US-0258928P.
PR	02-JAN-2001; 2001US-0259415P.
PR	04-JAN-2001; 2001US-0259785P.
PR	20-FEB-2001; 2001US-0269814P.
PR	09-MAR-2001; 2001US-0279863P.
PR	29-MAR-2001; 2001US-0279832P.
PR	29-MAR-2001; 2001US-0279833P.
PR	13-APR-2001; 2001US-0283889P.
PR	18-APR-2001; 2001US-0284447P.
PR	25-APR-2001; 2001US-0286683P.
PR	29-MAY-2001; 2001US-0294080P.
PR	16-AUG-2001; 2001US-0312915P.
PR	17-AUG-2001; 2001US-0313325P.
PR	17-SEP-2001; 2001US-0322699P.
PR	26-NOV-2001; 2001US-0333350P.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
PI	Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI	Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI	Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI	Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI	Gunther E, Smithson G, Millet I, Macdougall JR;
XX	
DR	WPI; 2002-732706/79.
DR	P-PSDB; ABP53589.
XX	
PT	New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT	associated disorders, such as cancers, neurological disorders, disorders
PT	of vesicular transport, gastrointestinal disorders, and autoimmune
XX	diseases.
PS	Claim 8; Page 123-125; 44pp; English.
XX	
CC	The present invention describes novel human proteins designated NOVX,
CC	where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC	cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC	antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC	immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC	protozoacide and antihelminthic activities, and can be used in gene
CC	therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC	manufacture of a medicament for treating a syndrome associated with a
CC	human disease selected from NOVX-associated disorder, such as cancers
CC	(e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC	ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC	ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC	disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC	diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC	disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC	autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC	anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC	and protozoal infections. The NOVX proteins can be used as immunogens to
CC	produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC	used in chromosome mapping, identifying individuals from minute
CC	biological samples (tissue typing), and in forensic identification of a
CC	biological sample. The present sequence encodes human NOV15d, which is
CC	located on chromosome 4
XX	
SQ	Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;
Alignment Scores:	
Pred. No.:	1.27e-72 Length: 8487
Score:	619.00 Matches: 111
Percent Similarity:	90.67% Conservative: 25
Best Local Similarity:	74.00% Mismatches: 14
Query Match:	79.16% Indels: 0
DB:	6 Gaps: 0
US-10-029-020-14_COPY_1250_1400 (1-151) x ABQ82346 (1-8487)	
QY	1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 3596 TACGTGCGCGGATATTCCTTCTGGAAATGTAAACAAGTGTCTTAGAACAATAAGAAATAAA 3655
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3656 GATTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCAGTCACG 3715
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 3716 GGAGATCTGTACGTTCTGACACAAACACCCGCAGAAATTATCGCCCAAGTCACCTTAGC 3775
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 3776 GGGCAAAAGACTTGACTAAAATCCAGAACTCGTCGCAGGACAGGGGAGCAATGCCTT 3835
QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 3836 CCGTTTGACGAGCGAGATGTGGGATGGAGGAAGGCCGTGGAAAGCCACACTCATGAGT 3895
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 3896 CCCAAAGGAATGCGAGTTGATAAGATGGATTAACTACTTTGTGATGGAACCATGATT 3955
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 3956 AGGAAAGTTGACCAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTGTGACTTCA 4015
QY 141 AlaArgProLeuSerCysAspSerValMet 150
Db 4016 GCCAGACCTTTAACTTGTGACACCAGCATG 4045

RESULT 6
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.
XX
AC ABQ82344;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15b encoding cDNA SEQ ID NO:37.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disease; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
CDS 151..8316
FT /*tag= a
FT /product= "NOV15b"
XX

PN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.

PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53587.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 114-117; 444pp; English.
PS
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15b, which is
CC located on chromosome 4
XX
SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:
Pred. NO.: 1.31e-72 Length: 8645
Score: 619.00 Matches: 111
Percent Similarity: 90.67% Conservative: 25
Best Local Similarity: 74.00% Mismatches: 14
Query Match: 79.16% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ABQ82344 (1-8645)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3772 TACGTGCGCGGATATTCCTTCTGGAAATGTAAACAAGTGTCTTAGAACAATAAGAAATAAA 3831
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3832 GATTTAGACATAGCAGCAACCCAGCTCATAGATACTACTTGTCAACGGATCCAGTCACG 3891
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 3892 GGAGATCTGTACGTTTCTGTGACACAAACACCCGCAGAAATTTATCGCCCAAGTCACTTACG 3951

Qy 61 ValVallYsAspLeuValLYsAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80

Db 3952 GGGGCAAAAGACTTGACTAAATGCAGAAGTCTCGCAGGGACAGGGGAGCAATGCCTT 4011

Qy 81 ProPheAspThrArgCysGlyAspGlyGlyLYsAlaThrGluAlaThrLeuThrAsn 100

Db 4012 CCGTTTGACGAGCGGAGATGTGGGGATGGAGGAAGCCGCTGGAAGCCACACTCATGAGT 4071

Qy 101 ProArgGlyIleThrValAspLYsPheGlyLeulleTyrPheValAspGlyThrMetIle 120

Db 4072 CCCAAAGGAATGGCAGTTGATAAGAATGGATTAAATCTACTTTGTTGATGGAACCATGATT 4131

Qy 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140

Db 4132 AGGAAAGTTGACCAAAATGGAATCATATCACTCTTCTGGGCTCTAACGATTGACTTCA 4191

Qy 141 AlaArgProLeuSerCysAspSerValMet 150

Db 4192 GCCAGACCTTTAACTTGTGACACCAGCATG 4221

RESULT 7

ABQ82343

ID ABQ82343 standard; cDNA; 8675 BP.

XX AC ABQ82343;

XX 17-DEC-2002 (first entry)

XX Human NOV15a encoding cDNA SEQ ID NO:35.

KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular; cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide; antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide; antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy; cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy; stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy; Pick's disease; vesicular transport disease; cystic fibrosis; goitre; diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine; ulcerative colitis; gastric reaction; duodenal disorder; infection; autoimmune disease; allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis; gene; chromosome 4; SS.

XX OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 151..8328

FT /*tag= a

FT /product= "NOV15a"

FT /transl_except= (pos:1249..1251,aa:Ser)

XX

PN WO200262999-A2.

XX

PD 15-AUG-2002.

XX

PF 31-DEC-2001; 2001WO-US049976.

XX

PR 29-DEC-2000; 2000US-0258928P.

PR 02-JAN-2001; 2001US-0259415P.

PR 04-JAN-2001; 2001US-0259785P.

PR 20-FEB-2001; 2001US-0269814P.

PR 09-MAR-2001; 2001US-0279863P.

PR 29-MAR-2001; 2001US-0279832P.

PR 29-MAR-2001; 2001US-0279833P.

PR 13-APR-2001; 2001US-0283889P.

PR 18-APR-2001; 2001US-0284447P.

PR 25-APR-2001; 2001US-0286683P.

PR 29-MAY-2001; 2001US-0294080P.

PR 16-AUG-2001; 2001US-0312915P.

PR 17-AUG-2001; 2001US-0313325P.

PR 17-SEP-2001; 2001US-0322699P.

PR 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

PA

XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

PI Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.

DR P-PSDB; ABP53586.

XX

PT New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.

PT

PT

XX Claim 8; Page 110-112; 444pp; English.

PS

XX The present invention describes novel human proteins designated NOVX, where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective, cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular, antiidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide, immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide, protozoacide and antihelminthic activities, and can be used in gene therapy. The NOVX proteins, nucleotides or antibodies can be used in the manufacture of a medicament for treating a syndrome associated with a human disease selected from NOVX-associated disorder, such as cancers (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle, ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke, ischaemic cerebrovascular disease, Alzheimer's disease or Pick's disease), disorders of vesicular transport (e.g. cystic fibrosis, diabetes mellitus, Grave's disease, or goitre), gastrointestinal disorders (e.g. ulcerative colitis, or gastric and duodenal disorders), autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic and protozoal infections. The NOVX proteins can be used as immunogens to produce antibodies and as vaccines. The NOVX nucleotide sequences may be used in chromosome mapping, identifying individuals from minute biological samples (tissue typing), and in forensic identification of a biological sample. The present sequence encodes human NOV15a, which is located on chromosome 4

XX

SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.31e-72 Length: 8675

Score: 619.00 Matches: 111

Percent Similarity: 90.67% Conservative: 25

Best Local Similarity: 74.00% Mismatches: 14

Query Match: 79.16% Indels: 0

DB: 6 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ABQ82343 (1-8675)

Qy 1 TyrTleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 3784 TACGTGGCGGGATATTCCCTTCTGAAATGTAAAGTGTCTAGAACTAAGAAATAAA 3843

Qy 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40

Db 3844 GATTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCACGGATCCAGTCACG 3903

Qy 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 3904 GGAGATCTGTACGTTTCTGACACAAACACCCGCAGAAATTATCGCCCAAGTCACTTACG 3963

Qy 61 ValVallYsAspLeuValLYsAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80

Db 3964 GGGGCAAAAGACTTGACTAAATGCAGAAGTCGTCGACGGGACAGGGGAGCAATGCCTT 4023

Qy 81 PropheAspAspThrArgCysGlyAspGlyGlyLYsAlaThrGluAlaThrLeuThrAsn 100

XX PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.

XX CC The invention relates to isolated polynucleotides which are

CC differentially-regulated in breast cancer. The methods and compositions

CC of the present invention are useful for detecting, diagnosing, staging,

CC monitoring, prognosticating, preventing and/or treating diseases and

CC conditions relating to breast cancer, and may be used in gene therapy or

CC antisense therapy. They can also be used in research, drug discovery,

CC clinical medicine and forensic medicine. Sequences given in records

CC ACC72012-ACC72074 represent polynucleotides of the invention that are

CC differentially-regulated in breast cancer. NOTE: The sequence data for

CC this patent did not form part of the printed specification, but was

CC obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.01e-71 Length: 9695

Score: 613.00 Matches: 112

Percent Similarity: 89.40% Conservative: 23

Best Local Similarity: 74.17% Mismatches: 16

Query Match: 78.39% Indels: 0

DB: 7 Gaps: 0

US-10-029-020-14_COPY_1250_1400 (1-151) x ACC72052 (1-9695)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20

Db 3783 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA 3842

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40

Db 3843 GAGTTAAACATAGCACAAACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTGTC 3902

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60

Db 3903 GGCTCGCTCTACGTGTCGACACCAACAGCAGGAGAATCTACCGGTCAAGTCTCTGAGT 3962

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80

Db 3963 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGAGCGGAGAGCAGTGTCCTA 4022

QY 81 ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100

Db 4023 CCCTTTGATGAAGCCCGCTGCGGGGATGGAGGAGGCCATAGATGCAACCCCTGATGAGC 4082

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120

Db 4083 CCGAGAGGTATTGCAGTAGACAGAATGGGCTCATGTACTTTGTCTGATGCCACCATGATC 4142

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140

Db 4143 CGGAAGGTGACCAGAAATGGAATCATCTCCACCCCTGCTGGGCTCCAATGACCTCACTGCC 4202

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151

Db 4203 GTCCGGCCGCTGAGCTGTGATTCCAGCATGGAT 4235

RESULT 10

AAS14089

ID AAS14089 standard; DNA; 9729 BP.

XX AAS14089;

AC AAS14089;

DT 18-DEC-2001 (first entry)

XX Human FCTR3f DNA sequence.

DE Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;

XX astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;

KW neurological disorder; neurodegenerative disorders; nerve trauma;

familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;

demyelinating Gardner syndrome; familial myelodysplastic syndrome;

mental health condition; immunological disorder; allergy; infertility;

bronchial asthma; Avellino type eosinophilia; lung disease; deafness;

reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;

desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;

gastric disorders; pancreatic disease; Schistosoma mansoni infection;

Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;

Corneal dystrophy-Greonow type I; Corneal dystrophy-lattice type I;

Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;

anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;

antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;

gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;

gene therapy; FCTR3a; neurestin-like protein; FCTR3f.

XX Homo sapiens.

OS

XX

FH Key Location/Qualifiers

FT 5'UTR 1..209

FT /*tag= b

FT CDS 210..8384

FT /*tag= a

FT /product= "Human FCTR3f"

FT 3'UTR 8385..9729

FT /*tag= c

XX

PN WO200166747-A2.

XX

PD 13-SEP-2001.

XX

PF 05-MAR-2001; 2001WO-US007160.

XX

PR 03-MAR-2000; 2000US-0186592P.

PR 03-MAR-2000; 2000US-0186718P.

PR 06-MAR-2000; 2000US-0187293P.

PR 06-MAR-2000; 2000US-0187294P.

PR 17-MAR-2000; 2000US-0190400P.

PR 07-APR-2000; 2000US-0196018P.

PR 03-JAN-2001; 2001US-0259548P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;

PI Macdougall J, Mishra V, Mezes PS, Rastelli L;

XX

DR WPI; 2001-596837/67.

DR P-PSDB; AAU08681.

XX

PT Novel polypeptides designated as FCTRX polypeptides, useful in detection,

PT prevention and treatment of a broad range of pathological states.

XX

PS Claim 9; Page 37-39; 215pp; English.

XX

CC The invention relates to human FCTRX polypeptides, FCTR1-FCTR7, and the

CC nucleic acids encoding them. These sequences are useful for the treatment

CC or prevention of numerous disorders including myelogenous leukaemia,

CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal

CC alloimmune thrombocytopaenia, neurological disorders, neurodegenerative

CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie

CC -Tooth neuropathy, demyelinating Gardner syndrome, familial

CC myelodysplastic syndrome, mental health conditions, immunological

CC disorders, allergy and infection, bronchial asthma, Avellino type

CC eosinophilia, lung diseases, reproductive disorders, infertility, male

CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,

CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric

CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni

CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,

CC Corneal dystrophy-Greonow type I, Corneal dystrophy-lattice type I and

CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding

CC the FCTR3a homologue FCTR3f

XX

SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:		1.02e-71	Length:	9729
Pred. No.:	613.00	Matches:	112	
Score:		Conservative:	23	
Percent Similarity:	89.40%	Mismatches:	16	
Best Local Similarity:	74.17%	Indels:	0	
Query Match:	78.39%	Gaps:	0	
DB:	5			
US-10-029-020-14_COPY_1250_1400 (1-151) x AAS14089 (1-9729)				
QY	1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20	
Db	3831	TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA	3890	
QY	21	AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer	40	
Db	3891	GAGTTTAAACATAGCAACACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTGTC	3950	
QY	41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60	
Db	3951	GGCTCGCTCTACGTGTCCGACACCAACAGCAGGAGAATCTACCGCTCAAGTCTCTGAGT	4010	
QY	61	ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu	80	
Db	4011	GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGGACGGAGAGCAGTGTCTA	4070	
QY	81	ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn	100	
Db	4071	CCCTTTGATGAAGCCCGCTGCGGGATGGAGGAAGGCCATAGATGCAACCTGATGAGC	4130	
QY	101	ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle	120	
Db	4131	CCGAGAGGTATTGCAGTAGACAAAGAATGGGCTCATGTACTTTGTGCGATGCCACCATGATC	4190	
QY	121	ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer	140	
Db	4191	CGGAAGGTTGACCAAGAATGGAATCATCTCACCCCTGCTGGGCTCCAATGACCTCACTGCC	4250	
QY	141	AlaArgProLeuSerCysAspSerValMetAsp	151	
Db	4251	GTCCGGCCGCTGAGCTGTGATTCCAGCATGGAT	4283	
RESULT 11				
ADB32028				
ID	ADB32028 standard; cDNA; 9729 BP.			
XX				
AC	ADB32028;			
XX				
DT	04-DEC-2003 (first entry)			
XX				
DE	Human FCTR3f cDNA.			
XX				
KW	Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;			
KW	myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;			
KW	ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;			
KW	renal cell carcinoma; melanoma; clear cell carcinoma;			
KW	granular cell carcinoma; neurological disorder;			
KW	neurodegenerative disorder; nerve trauma;			
KW	familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;			
KW	Gardner syndrome; mental health condition; immunological disorder;			
KW	allergy; asthma; lung disease; reproductive disorder; deafness;			
KW	glycoprotein deficiency; desmoid tumour; turcot syndrome;			
KW	liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;			
KW	diabetes; schistosoma mansoni infection; spinocerebellar ataxia;			
KW	plasmodium falciparum infection; Groenouw's corneal dystrophy;			
KW	lattice corneal dystrophy.			
XX				
OS	Homo sapiens.			
XX				
PN	US2003087816-A1.			
XX				
PD	08-MAY-2003.			
XX				

PF	05-MAR-2001; 2001US-00800198.			
XX				
PR	03-MAR-2000; 2000US-0186592P.			
XX				
PA	(VERM/)	VERMET C.		
PA	(FERN/)	FERNANDES E.		
PA	(SHIM/)	SHIMKETS R.		
PA	(HERR/)	HERRMANN J.		
PA	(MAJU/)	MAJUMDER K.		
PA	(MACD/)	MACDOUGALL J.		
PA	(MISH/)	MISHRA V.		
PA	(MEZE/)	MEZES P S.		
PA	(RAST/)	RASTELLI L.		
XX				
PPI	Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;			
PPI	Macdougall J, Mishra V, Mezes PS, Rastelli L;			
XX				
DR	WPI; 2003-625633/59.			
DR	P-PSDB; ADB32029.			
XX				
PPT	New FCTR3 polypeptide and encoding polynucleotide, useful for preventing			
PPT	or treating FCTR3-related disorders, such as cancer, autoimmune,			
PPT	neurodegenerative, gastrointestinal, reproductive and inflammatory			
PT	diseases.			
XX				
PS	Claim 9; Page 34-37; 155pp; English.			
XX				
CC	The invention relates to FCTR3 polypeptides and the polynucleotides			
CC	encoding them. The sequences of the invention are useful for the			
CC	manufacture of a medicament for diagnosing and treating disorders			
CC	associated with the FCTR3 polypeptide, such as colorectal cancer,			
CC	adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal			
CC	autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain			
CC	tumours, mammary tumours, human gliomas, astrocytomas, renal cell			
CC	carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear			
CC	cell and granular cell carcinomas, neurological disorders,			
CC	neurodegenerative disorders, nerve trauma, familial myelodysplastic			
CC	syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health			
CC	conditions, immunological disorders, allergy and infection, asthma, lung			
CC	diseases, male and female reproductive disorders, deafness, glycoprotein			
CC	deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis			
CC	C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni			
CC	infection, spinocerebellar ataxia, plasmodium falciparum infection,			
CC	Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence			
CC	represents cDNA encoding an FCTR3 polypeptide of the invention.			
XX				
SQ	Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;			
Alignment Scores:				
Pred. No.:	1.02e-71	Length:	9729	
Score:	613.00	Matches:	112	
Percent Similarity:	89.40%	Conservative:	23	
Best Local Similarity:	74.17%	Mismatches:	16	
Query Match:	78.39%	Indels:	0	
DB:	9	Gaps:	0	
US-10-029-020-14_COPY_1250_1400 (1-151) x ADB32028 (1-9729)				
QY	1	TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys	20	
Db	3831	TACATCCGACGCATCTTTCCCTCTCGAAATGTGACCAGCATCTTGGAGTTACGAAATAAA	3890	
QY	21	AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer	40	
Db	3891	GAGTTTAAACATAGCAACACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTGTC	3950	
QY	41	GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal	60	
Db	3951	GGCTCGCTCTACGTGTCCGACACCAACAGCAGGAGAATCTACCGCTCAAGTCTCTGAGT	4010	
QY	61	ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu	80	
Db	4011	GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGSCAGGGACGGAGCAGTGTCTA	4070	

ADB32023
ID ADB32023 standard; cDNA; 9826 BP.
XX AC ADB32023;
XX AC ADB32023;
DT 04-DEC-2003 (first entry)
XX Human FCTR3b cDNA.
DE Human; gene; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli;
XX myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX XX
OS Homo sapiens.
XX US2003087816-A1.
PN
XX
PD 08-MAY-2003.
XX
PF 05-MAR-2001; 2001US-00800198.
XX
PR 03-MAR-2000; 2000US-0186592P.
XX
PA (VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RAST/) RASTELLI L.
XX
PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2003-625633/59.
DR P-PSDB; ADB32024.
XX
PT New FCTRX polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTRX-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
XX
PS Claim 9; Page 29-32; 155pp; English.
XX
CC The invention relates to FCTRX polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTRX polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,

CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTRX polypeptide of the invention.
XX
SQ Sequence 9826 BP; 2567 A; 2645 C; 2476 G; 2138 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.03e-71 Length: 9826
Score: 613.00 Matches: 112
Percent Similarity: 89.40% Conservative: 23
Best Local Similarity: 74.17% Mismatches: 16
Query Match: 78.39% Indels: 0
DB: Gaps: 0
US-10-029-020-14_COPY_1250_1400 (1-151) x ADB32023 (1-9826)
QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db 3928 TACATCCGACGCATCTTCCCTCTCGAAATGTGACCAGCATCTTGAGATTACGAAATAAA 3987
QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspProMetSer 40
Db 3988 GAGTTTAAACATAGCAACAACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTGTCC 4047
QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db 4048 GGCTCGCTCTACGTGTCCGACACCAACAGCAGGAGAAATCTACCGCGTCAAGTCTCTGAGT 4107
QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db 4108 GGAACCAAGACCTGGTGGGAATTCGGAAGTTGTGGCAGGACGGGAGAGCAGTGTCTA 4167
QY 81 ProPheAspAspThrArgCysGlyAspGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db 4168 CCCTTTGATGAAGCCCGCTCGGGGATGGAGGGAAGGCCATAGATGCAACCCCTGATGAGC 4227
QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db 4228 CCGAGAGGTATTGCAGTAGACAAGAAATGGGCTCATGTACTTTGTCGATGCCACCATGATC 4287
QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db 4288 CGGAAGGTTGACCAGAATGGAATCATCTCCACCCTGCTGGGCTCCAATGACCTCACTGCC 4347
QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db 4348 GTCCGGCCCGCTGAGCTGTGATTCCAGCATGGAT 4380
RESULT 14
AAL60066
ID AAL60066 standard; cDNA; 4245 BP.
XX AC AAL60066;
XX 27-AUG-2003 (first entry)
XX Human Pc099 cDNA.
XX
KW Human; differentially regulated protein; prevention; therapy; vaccine;
KW prostate cancer; gene therapy; Pc099; teneurin-2; chromosome 5q34; gene;
KW ss.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
CDS 1..4056
FT /*tag= a
FT /product= "Human Pc099 protein"
FT /note= "No start codon"
FT /partial
XX
PN WO2003040331-A2.
XX 15-MAY-2003.
PD

```

XX PF 07-NOV-2002; 2002WO-US035563.
XX PF
XX PR 07-NOV-2001; 2001US-0331041P.
XX PR 07-NOV-2001; 2001US-0331042P.
XX PR 18-DEC-2001; 2001US-0340251P.
XX PR 07-JAN-2002; 2002US-0344791P.
XX PR
XX PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX PA
XX PI Sun Z, Li X, Jay G, Kovacs KF, Fan W;
XX PR
XX DR WPI; 2003-449451/42.
XX DR P-PSDB; AA029571.
XX PT New polynucleotide for diagnosing, staging, monitoring, prognosticating,
XX PT preventing or treating, or determining the predisposition to, diseases or
XX PT conditions such as prostate cancer, and for research or forensic science.
XX PT
XX PS Claim 29; Page 158-163; 100pp; English.
XX PS
XX CC The present invention relates to novel differentially regulated genes and
XX CC polypeptides encoded by them. Sequences of the invention are useful in
XX CC diagnosing, staging, monitoring, prognosticating, preventing, treating or
XX CC determining the predisposition to diseases or conditions such as prostate
XX CC cancer. They may be used as molecular markers, drug targets, vaccines, in
XX CC gene therapy, research, clinical medicine or forensic science. The
XX CC present sequence is a differentially regulated prostate cDNA, Pc099 which
XX CC codes for teneurin-2. Pc099 gene is located on chromosome 5q34
XX CC
XX SQ Sequence 4245 BP; 1035 A; 1132 C; 1115 G; 963 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.36e-66 Length: 4245
Score: 571.50 Matches: 108
Percent Similarity: 84.77% Conservative: 20
Best Local Similarity: 71.52% Mismatches: 16
Query Match: 73.08% Indels: 7
DB: 7 Gaps: 1

US-10-029-020-14_COPY_1250_1400 (1-151) x AAL60066 (1-4245)

QY 1 TyrIleArgArgIlePheProSerGlyAsnValThrAsnIleLeuGluLeuArgAsnLys 20
Db |||||||
Db 2365 TACATCCGACGCATCTTCCCTCTCGAATGTGACCAGCATCTTGGAGTTACGCAAC--- 2421

QY 21 AspPheArgHisSerHisSerProAlaHisLysTyrTyrLeuAlaThrAspPrometSer 40
Db ::|||
Db 2422 -----AACCCAGCACACAAGTACTACTTGGCAGTGGACCCCGTGTC 2463

QY 41 GlyAlaValPheLeuSerAspSerAsnSerArgArgValPheLysIleLysSerThrVal 60
Db |||:::
Db 2464 GGTCGCTCTACGTGTCGACACCAACAGCAGGAGAATCTACCGCGTCAAGTCTCTGAGT 2523

QY 61 ValValLysAspLeuValLysAsnSerGluValValAlaGlyThrGlyAspGlnCysLeu 80
Db |||||
Db 2524 GGAACCAAGACCTGGCTGGGAATTCGGAAGTTGTGGCAGGACGGGAGAGCAGTGTCTA 2583

QY 81 ProPheAspAspThrArgCysGlyAspGlyGlyLysAlaThrGluAlaThrLeuThrAsn 100
Db |||||:::
Db 2584 CCCTTTGATGAAGCCCGCTGCGGGATGGAGGAAGGCCCAATAGATGCAACCTGTATGAGC 2643

QY 101 ProArgGlyIleThrValAspLysPheGlyLeuIleTyrPheValAspGlyThrMetIle 120
Db |||||
Db 2644 CCGAGAGGTATTGCAGTAGACAAAGATGGGCTCATGTACTTTGTGGATGCCACCATGATC 2703

QY 121 ArgArgIleAspGlnAsnGlyIleIleSerThrLeuLeuGlySerAsnAspLeuThrSer 140
Db |||:::
Db 2704 CGGAGGTTGACCAAGATGGAATCATCTCCACCCCTGTGGGCTCCAATGACCTCACTGCC 2763

QY 141 AlaArgProLeuSerCysAspSerValMetAsp 151
Db |||||
Db 2764 GTCCGGCCGCTGAGCTGTGATTCCAGCATGGAT 2796
```

```

RESULT 15
ABK92230
ID ABK92230 standard; DNA; 12879 BP.
XX AC ABK92230;
XX DT 15-AUG-2002 (first entry)
XX DE Prostate cancer-associated DNA sequence #116.
XX KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
XX KW gene therapy; gene; ds.
XX OS Mammalia.
XX PN WO200230268-A2.
XX PD 18-APR-2002.
XX PF 12-OCT-2001; 2001WO-US032045.
XX PR 13-OCT-2000; 2000US-00687576.
XX PR 08-DEC-2000; 2000US-00733288.
XX PR 08-DEC-2000; 2000US-00733742.
XX PR 24-JAN-2001; 2001US-0263957P.
XX PR 16-MAR-2001; 2001US-0276791P.
XX PR 16-MAR-2001; 2001US-0276888P.
XX PR 06-APR-2001; 2001US-0281922P.
XX PR 24-APR-2001; 2001US-0286214P.
XX PR 30-APR-2001; 2001US-00847046.
XX PR 04-MAY-2001; 2001US-0288589P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX WPI; 2002-471335/50.
XX P-PSDB; ABG61913.

Detecting a prostate cancer-associated transcript in a cell in a patient,
useful for diagnosing prostate cancer (PC) or screening modulators of PC,
by determining if prostate cancer-associated genes are expressed in a
prostate tissue.

Claim 22; Page 394-397; 436pp; English.

The present invention relates to methods of detecting a prostate cancer-
associated transcript in a cell from a patient. The method comprises
contacting a biological sample from the patient with prostate cancer-
associated polynucleotides (designated PC genes) that selectively
hybridise to a sequence that is at least 80% identical to them. The
prostate cancer-associated polynucleotide sequences are differentially
expressed in prostate tumour tissue or in prostate cancer and are derived
from the tissues of various organisms such as humans or other mammals
(e.g. mice, sheep and dogs). The methods of the invention are useful for
diagnosing and treating prostate cancer in mammals. The prostate cancer-
associated genes are useful for diagnosing or treating prostate cancer,
as well as for identifying modulators of prostate cancer or agents that
inhibit prostate cancer. The nucleic acid sequences are particularly
useful in gene therapy, as a vaccine or in antisense applications.
ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
sequences
```

Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7.46e-61	Length:	12879
Score:	534.50	Matches:	100
Percent Similarity:	82.78%	Conservative:	25
Best Local Similarity:	66.23%	Mismatches:	19
Query Match:	68.35%	Indels:	7
DB:	6	Gaps:	1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 202.425 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200
Perfect score: 549
Sequence: 1 RLFRKWFAAAPDLSYFIWD.....IQSGILHKGNGENQFVSQQP 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020 @CGN 1 1 2156 @runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
†						

1	549	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
2	549	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
3	538.5	98.1	8355	13	US-10-383-201-55	Sequence 55, Appl
4	538.5	98.1	8438	13	US-10-042-865-1	Sequence 1, Appli
5	453	82.5	8473	17	US-10-038-854-39	Sequence 39, Appl
6	453	82.5	8487	17	US-10-038-854-41	Sequence 41, Appl
7	453	82.5	8645	17	US-10-038-854-37	Sequence 37, Appl
8	453	82.5	8675	17	US-10-038-854-35	Sequence 35, Appl
9	426	77.6	8409	9	US-09-808-602-79	Sequence 79, Appl
10	426	77.6	8409	10	US-09-800-198-67	Sequence 67, Appl
11	417	76.0	8797	9	US-09-808-602-74	Sequence 74, Appl
12	417	76.0	8797	9	US-09-808-602-77	Sequence 77, Appl
13	417	76.0	8797	10	US-09-800-198-62	Sequence 62, Appl
14	417	76.0	8797	10	US-09-800-198-65	Sequence 65, Appl
15	416	75.8	6560	9	US-09-808-602-76	Sequence 76, Appl
16	416	75.8	6560	10	US-09-800-198-64	Sequence 64, Appl
17	416	75.8	9058	16	US-10-144-194A-79	Sequence 79, Appl
18	416	75.8	9695	16	US-10-144-194A-81	Sequence 81, Appl
19	411	74.9	12880	16	US-10-295-027-927	Sequence 927, App
20	407	74.1	9729	9	US-09-808-602-12	Sequence 12, Appl
21	407	74.1	9729	10	US-09-800-198-12	Sequence 12, Appl
22	407	74.1	9826	9	US-09-808-602-7	Sequence 7, Appli
23	407	74.1	9826	10	US-09-800-198-7	Sequence 7, Appli
24	406	74.0	8689	9	US-09-808-602-78	Sequence 78, Appl
25	406	74.0	8689	10	US-09-800-198-66	Sequence 66, Appl
26	341	62.1	466	16	US-10-125-968-61	Sequence 61, Appl
27	323	58.8	8575	13	US-10-072-012-143	Sequence 143, App
28	221	40.3	534	13	US-10-027-632-49758	Sequence 49758, A
29	221	40.3	534	13	US-10-027-632-49759	Sequence 49759, A
30	221	40.3	534	16	US-10-027-632-49758	Sequence 49758, A
31	221	40.3	534	16	US-10-027-632-49759	Sequence 49759, A
32	125	22.8	310	15	US-10-029-386-16277	Sequence 16277, A
33	125	22.8	559	15	US-10-029-386-2577	Sequence 2577, Ap
34	75	13.7	71132	13	US-10-087-192-1867	Sequence 1867, Ap
35	73.5	13.4	3673778	15	US-10-312-841-2	Sequence 2, Appli
36	73	13.3	1908	15	US-10-040-862-9600	Sequence 9600, Ap
37	73	13.3	1908	16	US-10-057-475B-9600	Sequence 9600, Ap
38	73	13.3	1908	16	US-10-154-884B-9600	Sequence 9600, Ap
39	73	13.3	2516	13	US-09-795-651-6	Sequence 6, Appli
40	73	13.3	3023	9	US-09-900-236-4	Sequence 4, Appli
41	73	13.3	3023	15	US-10-066-443-4	Sequence 4, Appli
42	73	13.3	3023	15	US-10-256-889-4	Sequence 4, Appli
43	73	13.3	3023	16	US-10-439-799-4	Sequence 4, Appli
44	73	13.3	6303	13	US-10-342-887-1424	Sequence 1424, Ap
45	73	13.3	6303	13	US-10-172-118-1424	Sequence 1424, Ap

ALIGNMENTS

RESULT 1
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

```
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.:      1.53e-70      Length:      8354
Score:          549.00      Matches:      101
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             13          Gaps:          0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-383-201-43 (1-8354)

QY      1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
      |||
Db      3332 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCTCTATTATTTTGGGAC 3391

QY      21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
      |||
Db      3392 AAGACAGACGCTCTACAACCAGAAAGGTGTTTGGGCTTTCAGAAAGCCTTTGTTCCGTGGGT 3451

QY      41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
      |||
Db      3452 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAAACAACAGTGTGCAG 3511

QY      61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
      |||
Db      3512 GGCTATGAAATTGACGCGTCCAGCTTGAGGATGGAGCCTAGACAAACATCATGCCCTC 3571

QY      81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
      |||
Db      3572 AACATTCAAAGTGGTATCCTGCACAAAGGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAG 3631

QY      101 Pro 101
      |||
Db      3632 CCT 3634

RESULT 2
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
```

```
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.:      1.53e-70      Length:      8354
Score:          549.00      Matches:      101
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             13          Gaps:          0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-029-020-13 (1-8354)

QY      1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
      |||
Db      3332 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCTCTATTATTTTGGGAC 3391

QY      21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
      |||
Db      3392 AAGACAGACGCTCTACAACCAGAAAGGTGTTTGGGCTTTCAGAAAGCCTTTGTTCCGTGGGT 3451

QY      41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
      |||
Db      3452 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAAACAACAGTGTGCAG 3511

QY      61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
      |||
Db      3512 GGCTATGAAATTGACGCGTCCAGCTTGAGGATGGAGCCTAGACAAACATCATGCCCTC 3571

QY      81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
      |||
Db      3572 AACATTCAAAGTGGTATCCTGCACAAAGGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAG 3631

QY      101 Pro 101
      |||
Db      3632 CCT 3634

RESULT 3
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
```


; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 5.71e-69 Length: 8355
Score: 538.50 Matches: 101
Percent Similarity: 99.02% Conservative: 0
Best Local Similarity: 99.02% Mismatches: 0
Query Match: 98.09% Indels: 1
DB: 13 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-383-201-55 (1-8355)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||
Db 3301 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCTATTATTTCATTTGGGAC 3360

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||
Db 3361 AAGACAGACGCTCTACAACCAGAGGTGTTTGGGCTTTCAGAGCCCTTTGTTCCGIGGGT 3420

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
|||||
Db 3421 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAACACAGTGTGTCAG 3480

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
|||||
Db 3481 GGCTATGAAATTGATGCTCCAGCTTGAGGATGGAGCTTAGACAAACATCATGCCCTC 3540

QY 81 AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln 99
|||||
Db 3541 AACATTCAAAGTGGTGGCATCCTGCACAAAGGGAATGGGGAGAACAGTTTGTGTCTCAG 3600

QY 100 GlnPro 101
|||||
Db 3601 CAGCCT 3606

RESULT 4
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:

; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M

; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennnda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:

Pred. No.: 5.79e-69 Length: 8438
Score: 538.50 Matches: 101
Percent Similarity: 99.02% Conservative: 0
Best Local Similarity: 99.02% Mismatches: 0
Query Match: 98.09% Indels: 1
DB: 13 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-042-865-1 (1-8438)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||
Db 3311 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCTATTATTTCATTTGGGAC 3370

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||
Db 3371 AAGACAGACGCTCTACAACCAGAGGTGTTTGGGCTTTCAGAGCCCTTTGTTCCGTGGGT 3430

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
|||||
Db 3431 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAACACAGTGTGTCAG 3490

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
|||||
Db 3491 GGCTATGAAATTGACGCGTCCAAGCTTGGAGGATGGAGCTTAGACAAACATCATGCCCTC 3550

QY 81 AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln 99
|||||
Db 3551 AACATTCAAAGTGGTGGCATCCTGCACAAAGGGAATGGGGAGAACAGTTTGTGTCTCAG 3610

QY 100 GlnPro 101
|||||
Db 3611 CAGCCT 3616

RESULT 5
US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1


```

; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 8487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-41

Alignment Scores:
Pred. No.:          3.67e-56          Length:      8487
Score:              453.00           Matches:      78
Percent Similarity: 91.09%           Conservative: 14
Best Local Similarity: 77.23%         Mismatches:   9
Query Match:        82.51%           Indels:       0
DB:                 17              Gaps:         0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-038-854-41 (1-8487)
QY      1  ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db      3146 AGACTCTTCCAAAAGTGGTTTCCTGCCTCACCAAACTTGGCCTATACTTTTCATATGGGAT 3205

QY      21  LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db      3206 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGCAGTTGGA 3265

QY      41  TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db      3266 TATGAGTATGAGTCGTCGTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 3325

QY      61  GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisHisAlaLeu 80
Db      3326 GGCTATGAATTGGATGCGTCGTCACATGGGTGGCTGGACATTAGATAAACATCACGTCGTG 3385

QY      81  AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db      3386 GATGTACAGAACGGTATACTGTACAAAGGGAACCGGGGAAAACCCAGTTTCATCTCCCAGCAG 3445

QY      101 Pro 101
Db      3446 CCT 3448

RESULT 7
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
```

```

; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 37
; LENGTH: 8645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-37

Alignment Scores:
Pred. No.:          3.77e-56          Length:      8645
Score:              453.00           Matches:      78
Percent Similarity: 91.09%           Conservative: 14
Best Local Similarity: 77.23%         Mismatches:   9
Query Match:        82.51%           Indels:       0
DB:                 17              Gaps:         0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-038-854-37 (1-8645)
QY      1  ArgLeuPheArgLysTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db      3322 AGACTCTTCCAAAAGTGGTTTCCTGCCTCACCAAACTTGGCCTATACTTTTCATATGGGAT 3381

QY      21  LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db      3382 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGCAGTTGGA 3441

QY      41  TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db      3442 TATGAGTATGAGTCGTCGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 3501

QY      61  GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisHisAlaLeu 80
Db      3502 GGCTATGAATTGGATGCGTCGTCACATGGGTGGCTGGACATTAGATAAACATCACGTCGTG 3561

QY      81  AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db      3562 GATGTACAGAACGGTATACTGTACAAAGGGAACCGGGGAAAACCCAGTTTCATCTCCCAGCAG 3621

QY      101 Pro 101
Db      3622 CCT 3624
```


RESULT 8
US-10-038-854-35
; Sequence 35, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35

Alignment Scores:
Pred. No.: 3.79e-56 Length: 8675
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 17 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-10-038-854-35 (1-8675)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||:::||||| |||:::||||| |||:::||||| |||:::|||||
Db 3334 AGACTTTCAAAAGTGGTTTCCTGCCTCACCAAACCTGGCCTATACTTTTCATATGGGAT 3393
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||:::||||| |||:::||||| |||:::||||| |||:::|||||
Db 3394 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTCAGTTGGA 3453
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGln 60
|||||:::||||| |||:::||||| |||:::||||| |||:::|||||
Db 3454 TATGAGTATGAGTCGTGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 3513
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
|||||:::||||| |||:::||||| |||:::||||| |||:::|||||
Db 3514 GGCTATGAATTGGATGCGTCCCAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 3573
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
:::|||||:::||||| |||:::||||| |||:::||||| |||:::|||||
Db 3574 GATGTACAGAACGGTATACTGTACAGGGAAACGGGGAAACCAGTTTCATCTCCCAGCAG 3633
QY 101 Pro 101
|||
Db 3634 CCT 3636
RESULT 9
US-09-808-602-79
; Sequence 79, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Heriman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-808-602-79
Alignment Scores:
Pred. No.: 3.99e-52 Length: 8409
Score: 426.00 Matches: 74
Percent Similarity: 89.00% Conservative: 15
Best Local Similarity: 74.00% Mismatches: 11
Query Match: 77.60% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-79 (1-8409)

QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
|||||:::||||| |||:::||||| |||:::||||| |||:::|||||
Db 3409 CTATTTCAAAAATCAATTCTGGCATCTCCCAACITGGCTTATACATTTCATCTGGGACAAA 3468
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
|||||:::||||| |||:::||||| |||:::||||| |||:::|||||
Db 3469 ACAGATGCATATGGTCAAGAGTTTATGGGTTGTGAGATGCTGTAGTTTCTGTGGGTTT 3528

QY 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 3529 GAATATGAGACTTGTCCCACTTTGATTCTGTGGGAGAAAAGGACTGCGCTGCTGCAAGGA 3588
QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTyrTrpSerLeuAspLysHisHisAlaLeuAsn 81
: : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3589 TTTGAGCTAGATCCCTTCCAATCTAGGAGGATGGTCTTTGGATAAACATCATGTACTGAAT 3648
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
: : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3649 GTCAAGAGTGGTATATTGCACAAAGGCAATGGAGAAAATCAGTTTCTAACTCAGCAGCCA 3708

RESULT 10
US-09-800-198-67
; Sequence 67, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 67
; LENGTH: 8409
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-800-198-67

Alignment Scores:
Pred. No.: 3.99e-52 Length: 8409
Score: 426.00 Matches: 74
Percent Similarity: 89.00% Conservative: 15
Best Local Similarity: 74.00% Mismatches: 11
Query Match: 77.60% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-800-198-67 (1-8409)
QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 3409 CTATTCAAAAATCATTTCTGGCATCTCCCACTTGGCTTATACATTCATCTGGGACAAA 3468
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
| | | | | | | | | | | | | | | : : : | | | | | | | | | : : :
Db 3469 ACAGATGCATATGGTCAGAAGGTTTATGGGTTGTCAGATGCTGTAGTTTCTGTGGGTTT 3528
QY 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 3529 GAATATGAGACTTGTCCCACTTTGATTCTGTGGGAGAAAAGGACTGCGCTGCTGCAAGGA 3588
QY 62 TyrGluIleAspAlaSerLysLeuGlyTyrTrpSerLeuAspLysHisHisAlaLeuAsn 81
: : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3589 TTTGAGCTAGATCCCTTCCAATCTAGGAGGATGGTCTTTGGATAAACATCATGTACTGAAT 3648
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
: : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3649 GTCAAGAGTGGTATATTGCACAAAGGCAATGGAGAAAATCAGTTTCTAACTCAGCAGCCA 3708
RESULT 11
US-09-808-602-74
; Sequence 74, Application US/09808602

; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-808-602-74

Alignment Scores:
Pred. No.: 9.47e-51 Length: 8797
Score: 417.00 Matches: 72
Percent Similarity: 88.00% Conservative: 16
Best Local Similarity: 72.00% Mismatches: 12
Query Match: 75.96% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-74 (1-8797)
QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 3407 CTCCTCCAGAAGTCATTCCAGGCTTCTCCCAACCTAGCCTACACATTCATCTGGGACAAG 3466
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
| | | | | | | | | | | | | | | : : : | | | | | | | | | : : :
Db 3467 ACAGATGCTTATGGCCAAAGGGTTTATGGCCTATCGGATGCTGTGTGGGTTT 3526
QY 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
| | | | | : : : | | | | | | | | | | | | | | | : : : | | | | |
Db 3527 GAATATGAGACCTGCCCCAGTCTCATCTGTGGGAGAAAAGGACAGCCCTGTCTTCAGGGA 3586
QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTyrTrpSerLeuAspLysHisHisAlaLeuAsn 81
: : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3587 TTCAGCTGGACCCCTTCCCAACCTTGGAGGCTGGTCCCTGGACAAACACACCCCTCAAT 3646
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
: : : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3647 GTGAAAAGCGGAATACTACACAAAGGGACAGGGGAGAACCAAGTTCCTGACCCAGCAGCCT 3706
RESULT 12
US-09-808-602-77
; Sequence 77, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14

```
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 77
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-808-602-77

Alignment Scores:
Pred. No.:          9.47e-51          Length:      8797
Score:              417.00            Matches:      72
Percent Similarity: 88.00%            Conservative: 16
Best Local Similarity: 72.00%          Mismatches:   12
Query Match:        75.96%            Indels:        0
DB:                  9                Gaps:          0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-77 (1-8797)

QY      2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
Db      3407 CTCCTCCAGAAGTCATTCCAGGCTTCTCCCAACCTAGCCTACACATTCAATCTGGGACAAG 3466

QY      22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3467 ACAGATGCTTATGGCCAAAGGTTTATGGCCTATCGGATGCTGTGTCTGTGGGTTT 3526

QY      42 GluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGlnGly 61
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
Db      3527 GAATATGAGACCTGCCCCAGTCTCATCCTGTGGGAGAAAGGACAGCCCTGCTTCAGGGA 3586

QY      62 TyrGluIleAspAlaSerLysLeuGlyTyrTrpSerLeuAspLysHisAlaLeuAsn 81
      :::::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3587 TTCGAGCTGGACCCCTTCCAAACCTTGGAGGCTGGTCCCTGGACAAACACACCCCTCAAT 3646

QY      82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
      :::::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3647 GTGAAAAGCGGAATACTACAAAGGGACAGGGGAGAACCAAGTTCCTGACCCAGCAGCCT 3706

RESULT 13
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62

Alignment Scores:
Pred. No.:          9.47e-51          Length:      8797
Score:              417.00            Matches:      72
Percent Similarity: 88.00%            Conservative: 16
Best Local Similarity: 72.00%          Mismatches:   12
Query Match:        75.96%            Indels:        0
DB:                  9                Gaps:          0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-77 (1-8797)

QY      2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
Db      3407 CTCCTCCAGAAGTCATTCCAGGCTTCTCCCAACCTAGCCTACACATTCAATCTGGGACAAG 3466

QY      22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3467 ACAGATGCTTATGGCCAAAGGTTTATGGCCTATCGGATGCTGTGTCTGTGGGTTT 3526

QY      42 GluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGlnGly 61
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
Db      3527 GAATATGAGACCTGCCCCAGTCTCATCCTGTGGGAGAAAGGACAGCCCTGCTTCAGGGA 3586

QY      62 TyrGluIleAspAlaSerLysLeuGlyTyrTrpSerLeuAspLysHisAlaLeuAsn 81
      :::::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3587 TTCGAGCTGGACCCCTTCCAAACCTTGGAGGCTGGTCCCTGGACAAACACACCCCTCAAT 3646

QY      82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
      :::::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3647 GTGAAAAGCGGAATACTACAAAGGGACAGGGGAGAACCAAGTTCCTGACCCAGCAGCCT 3706

RESULT 13
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62

Alignment Scores:
Pred. No.:          9.47e-51          Length:      8797
Score:              417.00            Matches:      72
Percent Similarity: 88.00%            Conservative: 16
```

```
Best Local Similarity: 72.00%          Mismatches:   12
Query Match:          75.96%            Indels:        0
DB:                    10                Gaps:          0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-800-198-62 (1-8797)

QY      2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
Db      3407 CTCCTCCAGAAGTCATTCCAGGCTTCTCCCAACCTAGCCTACACATTCAATCTGGGACAAG 3466

QY      22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3467 ACAGATGCTTATGGCCAAAGGTTTATGGCCTATCGGATGCTGTGTCTGTGGGTTT 3526

QY      42 GluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGlnGly 61
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
Db      3527 GAATATGAGACCTGCCCCAGTCTCATCCTGTGGGAGAAAGGACAGCCCTGCTTCAGGGA 3586

QY      62 TyrGluIleAspAlaSerLysLeuGlyTyrTrpSerLeuAspLysHisAlaLeuAsn 81
      :::::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3587 TTCGAGCTGGACCCCTTCCAAACCTTGGAGGCTGGTCCCTGGACAAACACACCCCTCAAT 3646

QY      82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
      :::::||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3647 GTGAAAAGCGGAATACTACAAAGGGACAGGGGAGAACCAAGTTCCTGACCCAGCAGCCT 3706

RESULT 14
US-09-800-198-65
; Sequence 65, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-65

Alignment Scores:
Pred. No.:          9.47e-51          Length:      8797
Score:              417.00            Matches:      72
Percent Similarity: 88.00%            Conservative: 16
Best Local Similarity: 72.00%          Mismatches:   12
Query Match:        75.96%            Indels:        0
DB:                  10                Gaps:          0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-800-198-65 (1-8797)

QY      2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
Db      3407 CTCCTCCAGAAGTCATTCCAGGCTTCTCCCAACCTAGCCTACACATTCAATCTGGGACAAG 3466

QY      22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      3467 ACAGATGCTTATGGCCAAAGGTTTATGGCCTATCGGATGCTGTGTCTGTGGGTTT 3526

QY      42 GluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGlnGly 61
      |||||:::||||| |||||:::||||| |||||:::||||| |||||:::|||||
```


Db 3527 GAATATGAGACCTGCCCCAGTCTCATCCTGTGGGAGAAAAGGACAGCCCTGCTTCAGGGA 3586

Qy 62 TyrGluIleAspAlaSerLysLeuGlyTyrSerLeuAspLysHisHisAlaLeuAsn 81
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 3587 TTCGAGCTGGACCCCTTCCAACTTGGAGGCTGGTCCCTGGACAAACACACACCCCTCAAT 3646

Qy 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 3647 GTGAAAAGCGGAATACTACACAAAGGGACAGGGGAGAACCAAGTTCTTGACCCAGCAGCCT 3706

RESULT 15

US-09-808-602-76

; Sequence 76, Application US/09808602

; Patent No. US20020155115A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Corine A

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrman, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishnu

; APPLICANT: Mezes, Peter S

; APPLICANT: MacDougall, John

; TITLE OF INVENTION: No. US20020155115A1e1 Proteins and Nuclec Acids Encoding Same

; FILE REFERENCE: 15966-697 CIP

; CURRENT APPLICATION NUMBER: US/09/808,602

; CURRENT FILING DATE: 2001-03-14

; PRIOR APPLICATION NUMBER: 09/800,198

; PRIOR FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 76

; LENGTH: 6560

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-808-602-76

Alignment Scores:

Pred. No.:	8.82e-51	Length:	6560
Score:	416.00	Matches:	72
Percent Similarity:	88.00%	Conservative:	16
Best Local Similarity:	72.00%	Mismatches:	12
Query Match:	75.77%	Indels:	0
DB:	9	Gaps:	0

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-808-602-76 (1-6560)

Qy 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 215 CTCTCCAGAAAGTCATTCACGGCTTCTCCCAACCTGGCCCTACACCTTCATCTGGGACAAG 274

Qy 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 275 ACAGATGCGTATGGCCAAAGGTTGTATGGAATCTCAGATGCTGTGTGTCTGTCTCGGGTTT 334

Qy 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 335 GAATATGAGACCTGTCCCAAGTCTAATTCTCTGGGAGAAAAGGACAGCCCTCCTTCAGGGA 394

Qy 62 TyrGluIleAspAlaSerLysLeuGlyTyrSerLeuAspLysHisHisAlaLeuAsn 81
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 395 TTCGAGCTGGACCCCTCCAACCTCGGTGGTGGTCCCTAGACAAACACACATCCTCAAT 454

Qy 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 455 GTTAAAGTGAATCCTACACAAAGGCACTGGGGAAACCAGTTCTTGACCCAGCAGCCT 514

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 29.8807 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200
Perfect score: 549
Sequence: 1 RLFRKWFAAAPDLSYFIWD.....IQSGILHKGNGENQFVSQQP 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020 @CGN_1_1_258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	13.3	3023	4	US-09-203-453-4 Sequence 4, Appli
2	73	13.3	3023	4	US-09-900-236-4 Sequence 4, Appli
3	69.5	12.7	1875	2	US-08-683-743-3 Sequence 3, Appli
4	69.5	12.7	2367	4	US-09-489-039A-5660 Sequence 5660, Ap
5	69.5	12.7	3532	4	US-09-221-017B-799 Sequence 799, App
6	68.5	12.5	3852	1	US-08-306-546C-1 Sequence 1, Appli
7	68.5	12.5	3852	2	US-08-530-524A-1 Sequence 1, Appli
8	66.5	12.1	4403765	3	US-09-103-840A-2 Sequence 2, Appli
9	66.5	12.1	4403765	3	US-09-103-840A-2 Sequence 2, Appli
10	66	12.0	898	1	US-08-419-078-3 Sequence 3, Appli
11	66	12.0	898	1	US-08-726-883-3 Sequence 3, Appli
12	66	12.0	1322	1	US-08-419-078-1 Sequence 1, Appli

13	66	12.0	1322	1	US-08-726-883-1	Sequence 1, Appli
14	66	12.0	1323	4	US-09-023-655-55	Sequence 55, Appl
15	66	12.0	2936	4	US-09-976-594-1044	Sequence 1044, Ap
16	66	12.0	3293	1	US-08-030-096-1	Sequence 1, Appli
c	17	65.5	552	4	US-09-134-000C-2495	Sequence 2495, Ap
c	18	65.5	732	6	5273901-2	Patent No. 5273901
c	19	65.5	732	6	5482709-1	Patent No. 5482709
20	65.5	11.9	984	4	US-09-894-844-89	Sequence 89, Appl
21	65.5	11.9	1358	2	US-08-160-524A-3	Sequence 3, Appli
22	65.5	11.9	1559	2	US-08-160-524A-1	Sequence 1, Appli
23	65.5	11.9	1886	1	US-08-461-773-15	Sequence 15, Appl
24	65.5	11.9	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	25	65.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
c	26	64.5	523	2	US-08-508-786-8	Sequence 8, Appli
c	27	64.5	523	5	PCT-US96-12158-8	Sequence 8, Appli
c	28	64.5	574	2	US-08-508-786-7	Sequence 7, Appli
c	29	64.5	574	5	PCT-US96-12158-7	Sequence 7, Appli
c	30	64.5	722	2	US-08-508-786-6	Sequence 6, Appli
c	31	64.5	722	5	PCT-US96-12158-6	Sequence 6, Appli
c	32	64.5	1030	2	US-08-508-786-5	Sequence 5, Appli
c	33	64.5	1030	5	PCT-US96-12158-5	Sequence 5, Appli
c	34	64.5	1294	2	US-08-508-786-4	Sequence 4, Appli
c	35	64.5	1294	5	PCT-US96-12158-4	Sequence 4, Appli
c	36	64.5	1372	2	US-08-508-786-3	Sequence 3, Appli
c	37	64.5	1372	5	PCT-US96-12158-3	Sequence 3, Appli
c	38	64.5	1988	2	US-08-508-786-2	Sequence 2, Appli
c	39	64.5	1988	5	PCT-US96-12158-2	Sequence 2, Appli
c	40	64.5	2010	2	US-08-508-786-1	Sequence 1, Appli
c	41	64.5	2010	5	PCT-US96-12158-1	Sequence 1, Appli
c	42	64.5	3470	4	US-09-976-594-345	Sequence 345, App
43	64	11.7	2610	2	US-08-989-386-2	Sequence 2, Appli
44	64	11.7	3045	4	US-09-134-000C-2860	Sequence 2860, Ap
45	63.5	11.6	1724	3	US-08-993-260-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-203-453-4
; Sequence 4, Application US/09203453
; Patent No. 6426411
; GENERAL INFORMATION:
; APPLICANT: Spiegelman, Bruce M.; Puigserver, Pere; Wu, Zhidan and Adelmant, Guillaume
; TITLE OF INVENTION: PGC-1, A NOVEL BROWN FAT PPAR{SYMBOL 103 \f "Symbol"} COACTIVATOR
; FILE REFERENCE: DFN-023CP
; CURRENT APPLICATION NUMBER: US/09/203,453
; CURRENT FILING DATE: 1998-12-01
; EARLIER APPLICATION NUMBER: 09/086,912
; EARLIER FILING DATE: 1998-05-29
; EARLIER APPLICATION NUMBER: 60/048,107
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 3023
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)..(2482)
US-09-203-453-4

Alignment Scores:
Pred. No.: 2.47 Length: 3023
Score: 73.00 Matches: 29
Percent Similarity: 41.58% Conservative: 13
Best Local Similarity: 28.71% Mismatches: 39
Query Match: 13.30% Indels: 20
DB: 4 Gaps: 5

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-203-453-4 (1-3023)

Qy 3 PheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLysThr 22


```
Db      77  TTCAGGAGCTGGATGGCG-----TGGGACATGTGC 106
QY      23  AspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyrGlu 42
Db      107 AACCAGGACTCTGAGTCTGTATGGAGTGACATCGAGTGTGCTGCTCTGGTTGGTGAAGAC 166
QY      43  TyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGlnGlyTyr 62
Db      167 CAGCCTCTTTGCCCAGATCTTCCT-----GAACCTTGATCTTTCTGAACATAGATGTGAAC 220
QY      63  GluIleAspAla---SerLysLeuGlyGly-----TrpSerLeuAspLysHisAla 79
Db      221 GACTGGATACAGACAGCTTTCTGGGTGGACTCAAGTGGTGCAGTGACCAATCAGAAATA 280
QY      80  LeuAsnIleGln-----SerGlyIleLeuHisLysGlyAsnGlyGluAsn 94
Db      281 ATATCCAATCAGTACACAACATGAGCCTTCAAACATATTTGAGAAGATAGATGAAGAGAAT 340
QY      95  Gln 95
Db      341 GAG 343

RESULT 2
US-09-900-236-4
; Sequence 4, Application US/09900236
; Patent No. 6525178
; GENERAL INFORMATION:
; APPLICANT: Spiegelman, Bruce M.; Puigserver, Pere; Wu, Zhidan and
; APPLICANT: Adelmant, Guillaume
; TITLE OF INVENTION: PGC-1, A NOVEL BROWN FAT PPAR{SYMBOL 103 \f "Symbol"}
; TITLE OF INVENTION: COACTIVATOR
; FILE REFERENCE: DFN-023CP
; CURRENT APPLICATION NUMBER: US/09/900,236
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/203,453
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-01
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/048,107
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 3023
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)..(2482)
US-09-900-236-4

Alignment Scores:
Pred. No.:      2.47      Length:      3023
Score:          73.00      Matches:      29
Percent Similarity: 41.58%      Conservative: 13
Best Local Similarity: 28.71%      Mismatches:  39
Query Match:     13.30%      Indels:      20
DB:              4        Gaps:         5

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-900-236-4 (1-3023)

QY      3  PheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLysThr 22
Db      77  TTCAGGAGCTGGATGGCG-----TGGGACATGTGC 106
QY      23  AspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyrGlu 42
Db      107 AACCAGGACTCTGAGTCTGTATGGAGTGACATCGAGTGTGCTGCTCTGGTTGGTGAAGAC 166
QY      43  TyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrValLeuGlnGlyTyr 62
Db      167 CAGCCTCTTTGCCCAGATCTTCCT-----GAACCTTGATCTTTCTGAACATAGATGTGAAC 220
QY      63  GluIleAspAla---SerLysLeuGlyGly-----TrpSerLeuAspLysHisAla 79
```

```
Db      221 GACTTGGATACAGACAGCTTTCTGGGTGGACTCAAGTGGTGCAGTGAACCAATCAGAAATA 280
QY      80  LeuAsnIleGln-----SerGlyIleLeuHisLysGlyAsnGlyGluAsn 94
Db      281 ATATCCAATCAGTACACAACATGAGCCTTCAAACATATTTGAGAAGATAGATGAAGAGAAT 340
QY      95  Gln 95
Db      341 GAG 343

RESULT 3
US-08-683-743-3
; Sequence 3, Application US/08683743
; Patent No. 5843697
; GENERAL INFORMATION:
; APPLICANT: Pestka, Sidney
; APPLICANT: Kotenko, Serguei
; TITLE OF INVENTION: CYTOKINE RECEPTOR SIGNAL TRANSDUCTION
; TITLE OF INVENTION: CHAIN
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,743
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 601-1-050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
US-08-683-743-3

Alignment Scores:
Pred. No.:      4.02      Length:      1875
Score:          69.50      Matches:      20
Percent Similarity: 45.21%      Conservative: 13
Best Local Similarity: 27.40%      Mismatches:  31
Query Match:     12.66%      Indels:       9
DB:              2        Gaps:         3

US-10-029-020-14_COPY_1100_1200 (1-101) x US-08-683-743-3 (1-1875)

QY      10  AlaProAspLeu---SerTyrTyrPheIleTrpAspLysThrAspValTyrAsnGlnLys 28
Db      440 GCCCCTAAATTGAGAATGAATACGAAACTTGGACTATGAAGAATGTGTATAACTCATGG 499
QY      29  ValPhe-----GlyLeuSerGluAlaPhe---ValSerValGly 40
Db      500 ACTTATAATGTGCAATACTTGAAAAACGGTACTGATGAAAAAGTTTCAAAATTACTCCCCAG 559
```

```
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 560 TATGACTTTGAGTCTCAGAAACCTGGAGCCATGGACAACTTATTGTGTTCAAGTTCGA 619
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSer 73
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 620 GGGTTTCTTCTGATCGGAACAAAGCTGGGAATGGAGT 658
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
US-09-489-039A-5660
; Sequence 5660, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5660
; LENGTH: 2367
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5660

Alignment Scores:
Pred. No.: 5.63 Length: 2367
Score: 69.50 Matches: 26
Percent Similarity: 36.97% Conservative: 18
Best Local Similarity: 21.85% Mismatches: 36
Query Match: 12.66% Indels: 39
DB: 4 Gaps: 6

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-489-039A-5660 (1-2367)

QY 3 PheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLysThr 22
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 676 TATCGCCGCTGGCAGGCGGAAGCCGGGACATCGATTATTACCTGTTTACC GGTAAG-- 732
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 23 AspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheVal----- 37
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 733 -----CAGGTGCTGGATGTGACCAAGCCTTCGTTTCGCTGACCGGTAAG 777
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 38 -----SerValGlyTyr-----GluTyrGluSer 45
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 778 ACGCTGTTTGGACCCAAATGGAGCCTCGGCTACAGCGGCTCCACCATGCACTATACCGAC 837
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 46 CysProAsp-----LeuIleLeuTrpGluLysArgThrThr 57
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 838 GCGCCGGATGCGCAAAATCAGCTGATGAACCTTATCCGCCCTGTGCGAGCAGCAGCGATC 897
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 58 ValLeuGlnGlyTyrGluIleAspAla-----SerLysLeuGlyGly----- 71
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 898 CCCTGCGATTCTTCAGCTCTCTCCGCTATACCTCAATCAACGGAAGCGGTACGTC 957
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 72 -----TrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIleLeuHis 88
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 958 TTTAACTGGAACCTACGACAAGGTGCGCGAGCCCAAGATGATGAGCCAGTCGTTTCAC 1014
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
US-09-221-017B-799/c
; Sequence 799, Application US/09221017B
; Patent No. 644799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
```

```
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 799:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3532 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...3532
US-09-221-017B-799

Alignment Scores:
Pred. No.: 10 Length: 3532
Score: 69.50 Matches: 23
Percent Similarity: 37.89% Conservative: 13
Best Local Similarity: 24.21% Mismatches: 32
Query Match: 12.66% Indels: 27
DB: 4 Gaps: 3

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-221-017B-799 (1-3532)

QY 32 LeuSerGluAlaPheValSerValGlyTyr-----GluTyrGlu 44
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2705 CTCGATCAATTTTTCGATGAGAAAGGCTACTACCCCGCATTCACAGGACCAAGGAA 2646
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 45 SerCysProAspLeuIleLeuTrpGluLysArgThrThrVal----- 58
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2645 ACCGACCCTACCATACACTCTTGGGAGAGCGGACGAGGCTAGCCGTTACGCTTGCTATT 2586
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 59 -----LeuGlnGlyTyrGluIleAspAla 66
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2585 GCTTACGATTATTGGTGCCTTAGCCGATTGGCACAGTTCTCCGGCCACGAGCAGGAGCG 2526
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 67 SerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


Db 1686 AGTGACTGCTGGAATGGAGAACTTTCCCAAGCTCGA---TTTGTATCT---GAGTAT 1739

QY 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
|||:::|||||
Db 1740 GGATATCAGTCTGGCCTTCTTCAGTACATTAGAAAAGGTTTCCTCT----- 1787

QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeuAsn 81
|||:::|||||
Db 1788 ---GAAGAGGACTGCTTACAGAAGCAGCTTTCACCTTCATCGGAACATTGATTAAAC 1844

QY 82 IleGlnSerGlyIleLeuHisLys 89
:::|||||:::
Db 1845 GGTAACAATGAATGCTTCACCAG 1868

RESULT 8
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 6.2e+05 Length: 4403765
Score: 66.50 Matches: 22
Percent Similarity: 36.27% Conservative: 15
Best Local Similarity: 21.57% Mismatches: 34
Query Match: 12.11% Indels: 31
DB: 3 Gaps: 4

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-103-840A-2 (1-4403765)

QY 4 ArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp----- 20
|||:::|||||
Db 3372455 CGGCTGTGGGTAGCA-----GACCTCACCTATGTGTCGACCTGGCAGGGTTCGCC 3372505

QY 21 -----LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPhe 36
|||||:::|||||
Db 3372506 TACGTGGCCTTTGTACCGCAGCGCTAGCTCGCAGGATCCTGGGC----- 3372550

QY 37 ValSerValGlyTyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThr 56
|||||:::|||||
Db 3372551 -----TGGCGGGTGCCTTCC 3372565

QY 57 ThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLys 76
|||:::|||||
Db 3372566 ACGATGGCCACCTCCATGGTCTCGACGCGATCGAGCAAGCCATCTGGACCCGCCAACAA 3372625

QY 77 HisHisAlaLeuAsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPhe 96
|||:::|||||
Db 3372626 GAAAGCGTACTCGACCTGAAAGACGTTATCCACCATACGGATAGGGGA---TCTCAGTAC 3372682

QY 97 ValSer 98
|||
Db 3372683 ACATCG 3372688

RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 6.2e+05 Length: 4403765
Score: 66.50 Matches: 22
Percent Similarity: 36.27% Conservative: 15
Best Local Similarity: 21.57% Mismatches: 34
Query Match: 12.11% Indels: 31
DB: 3 Gaps: 4

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-103-840A-2 (1-4403765)

QY 4 ArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp----- 20
|||:::|||||
Db 1986054 CGGCTGTGGGTAGCA-----GACCTCACCTATGTGTCGACCTGGGCGAGGGTTCGCC 1986004

QY 21 -----LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPhe 36
|||||:::|||||
Db 1986003 TACGTGGCCTTTGTACCGCAGCGCTACGCTCGCAGGATCCTGGGC----- 1985959

QY 37 ValSerValGlyTyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThr 56
|||||:::|||||
Db 1985958 -----TGGCGGGTGCCTTCC 1985944

QY 57 ThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLys 76
|||:::|||||
Db 1985943 ACGATGGCCACCTCCATGGTCTCGACGCGATCGAGCAAGCCATCTGGACCCGCCAACAA 1985884

QY 77 HisHisAlaLeuAsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPhe 96
|||:::|||||
Db 1985883 GAAAGCGTACTCGACCTGAAAGACGTTATCCACCATACGGATAGGGGA---TCTCAGTAC 1985827

QY 97 ValSer 98
|||
Db 1985826 ACATCG 1985821

RESULT 10
US-08-419-078-3
; Sequence 3, Application US/08419078
; Patent No. 5587306
; GENERAL INFORMATION:
; APPLICANT: HAWKINS, PHILLIP R.
; APPLICANT: SEILHAMER, JEFFREY J.
; TITLE OF INVENTION: PHOSPHOLIPASE C HOMOLOG
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3330 HILLVIEW AVENUE


```
Pred. No.: 1323
Score: 7.85
Length: 14
Percent Similarity: 66.00
Matches: 5
Best Local Similarity: 54.29%
Conservative: 12
Mismatch: 4
Indels: 1
Query Match: 12.02%
Gaps: 4
DB: 4

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-023-655-55 (1-1323)

QY 52 TrpGluLysArgThrThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 193 TGAAGAAAGCAGCACTGCATTAAA-----GATGTTGTCAAAGTTGGTGCA 240

QY 72 TrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 241 GTTGATGCAGATAAGCATCATTCCTAGGAGGTCAAGTATGGTGT 285

RESULT 15
US-09-976-594-1044
; Sequence 1044, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 1044
; LENGTH: 2936
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 1383263.14
US-09-976-594-1044

Alignment Scores:
Pred. No.: 24.9 Length: 2936
Score: 66.00 Matches: 14
Percent Similarity: 54.29% Conservative: 5
Best Local Similarity: 40.00% Mismatches: 12
Query Match: 12.02% Indels: 4
DB: 4

US-10-029-020-14_COPY_1100_1200 (1-101) x US-09-976-594-1044 (1-2936)

QY 52 TrpGluLysArgThrThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGly 71
Db 299 TGAAGAAAGCAGCACTGCATTAAA-----GATGTTGTCAAAGTTGGTGCA 346

QY 72 TrpSerLeuAspLysHisAlaLeuAsnIleGlnSerGlyIle 86
Db 347 GTTGATGCAGATAAGCATCATTCCTAGGAGGTCAAGTATGGTGT 391

Search completed: August 14, 2004, 20:09:55
Job time : 1824.88 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 1184.51 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200
Perfect score: 549
Sequence: 1 RLFRKWFAPLSDLYFYFWD.....IQSGILHKGNGENQFVSQQP 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=20000000000
-USER=US10029020_@CGN_1_13135_@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	549	100.0	6246	29	AY413475	AY413475 Homo sapi
2	544	99.1	3190	29	AY413477	AY413477 Mus muscu
3	544	99.1	5970	29	AY413476	AY413476 Pan trogl
4	454	82.7	5087	29	AY405422	AY405422 Mus muscu
5	453	82.5	5069	29	AY405421	AY405421 Pan trogl
6	453	82.5	5094	29	AY405420	AY405420 Homo sapi
7	422	76.9	796	14	CD578516	CD578516 UI-M-GIO-
8	403	73.4	750	9	AL710250	AL710250 DKFZp686F
9	391	71.2	539	9	AL919858	AL919858 AL919858
10	390	71.0	725	14	CA752216	CA752216 UI-M-FOO-
11	384	69.9	789	14	CD803315	CD803315 UI-M-GV0-
12	376	68.5	348	14	CB704746	CB704746 AMGNNUC:S
13	348	63.4	262	10	BB538418	BB538418 BB538418
14	290	52.8	318	10	BB501284	BB501284 BB501284
15	287	52.3	505	12	BJ016612	BJ016612 BJ016612
16	268	48.8	728	14	CF536185	CF536185 UI-M-GIO-
17	248.5	45.3	919	29	CNS01K9E	AL148051 Anopheles
18	223	40.6	692	14	CF530821	CF530821 UI-M-GH0-
19	223	40.6	706	29	CE611069	CE611069 tigr-gss-
20	222	40.4	529	9	AL699161	AL699161 DKFZp686F
21	218	39.7	716	29	AG098731	AG098731 Pan trogl
22	215.5	39.3	936	29	CNS022KC	AL221061 Tetraodon
23	210	38.3	833	29	CNS02D8V	AL192136 Tetraodon
24	208.5	38.0	346	28	AQ970978	AQ970978 RPCI-23-3
25	203	37.0	778	13	BU705631	BU705631 UI-M-FOO-
26	201	36.6	736	14	CD803545	CD803545 UI-M-GV0-
27	170	31.0	730	14	CF182768	CF182768 UI-M-EY0-
28	161.5	29.4	1146	28	CC238800	CC238800 CH261-50M
29	156	28.4	835	28	BZ165653	BZ165653 CH230-463
30	141	25.7	851	29	CNS03J5I	AL246447 Tetraodon
31	141	25.7	990	29	CNS0428B	AL271172 Tetraodon
32	134	24.4	975	29	CNS05RSP	AL350962 Tetraodon
33	132	24.0	619	29	CI22E8	AJ226991 Ciona int
34	124	22.6	480	9	AI387033	AI387033 GH17448.5
35	123	22.4	735	14	CF745263	CF745263 UI-M-GV0-
36	122.5	22.3	968	29	CNS00765	AL066865 Drosophil
37	107	19.5	726	29	CC574461	CC574461 CH240_451
38	106	19.3	1052	29	CNS05D17	AL331828 Tetraodon
39	106	19.3	1085	29	CNS0533P	AL318958 Tetraodon
40	105	19.1	1201	29	CNS0109M	AL098692 Drosophil
41	100	18.2	916	29	CNS04M38	AL296909 Tetraodon
42	94	17.1	243	28	AZ647907	AZ647907 IM0514K07
43	84	15.3	645	28	BH073129	BH073129 RPCI-24-2
44	84	15.3	726	29	CC611675	CC611675 OGUAC85TM
45	84	15.3	733	29	CE320811	CE320811 tigr-gss-

ALIGNMENTS

RESULT 1
AY413475
LOCUS AY413475 6246 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY413475
VERSION AY413475.1 GI:39769437
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 6246)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 6246)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..6246
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>6246
/locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.: 1.38e-63 Length: 6246
Score: 549.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY413475 (1-6246)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 1249 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCTATTATTTCATTTGGGAC 1308
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 1309 AAGACAGACGCTCTACAACCAGAAGGTGTTGGGCTTCAGAAAGCCTTTGTTCCGTGGGT 1368
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 1369 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGTCTGCAG 1428
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 1429 GGCTATGAATTTGACGGCTCCAAGCTTGAGGATGGAGCGCTAGACAAACATCATGCCCCTC 1488
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 1489 AACATTCAAAGTGGCATCCTGCACAAAGGAATGGGGAGAACCAAGTTTGTGTCTCAGCAG 1548
QY 101 Pro 101
Db 1549 CCT 1551
RESULT 2
AY413477 3190 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY413477
VERSION AY413477.1 GI:39769439
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 3190)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3190)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES Location/Qualifiers
source 1..3190
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
gene <1..>3190
/locus_tag="HCM4903"
ORIGIN
Alignment Scores:
Pred. No.: 2.57e-63 Length: 3190
Score: 544.00 Matches: 100
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 1
Query Match: 99.09% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY413477 (1-3190)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 1249 CGGCTCTTCAGGAAGTGGTTTCCCGCAGCTCCTGACCTGTCTATTACTTCATCTGGGAC 1308
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 1309 AAGACAGATGTCTACAACCAGAAGGTGTTTGGACTCTCAGAAGCCTTTGTTCCGTGGGT 1368
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 1369 TATGAGTATGAATCCTGCCAGATCTGATCCTGTGGGAGAAAAGGACAGCGTCTGCAG 1428
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 1429 GGCTATGAATTTGATGCTCCAAGCTGGGGGCTGGAGTCTGGATAAGCACCATGCCCCTG 1488
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 1489 AACATCCAGAGTGGCATCCTGCACAAAGGAATGGAGAGAACCAAGTTTGTGTCTCCAGCAG 1548
QY 101 Pro 101
Db 1549 CCA 1551
RESULT 3
AY413476 5970 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY413476
VERSION AY413476.1 GI:39769438
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 5970)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5970)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source Location/Qualifiers
1..5970
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5970
/locus_tag="HCM4903"
gene
ORIGIN
Alignment Scores:
Pred. No.: 6.32e-63 Length: 5970
Score: 544.00 Matches: 100
Percent Similarity: 99.01% Conservative: 0
Best Local Similarity: 99.01% Mismatches: 1
Query Match: 99.09% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY413476 (1-5970)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 973 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGATCTGTCCCTATTATTTCATTTGGGAC 1032
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 1033 AAGACAGACGCTCTACAACCAAGAGGTGTTGGGCTTTCAGAGCCCTTGTTCCGTGGGT 1092
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 1093 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAACAGCAGTGTGCAG 1152
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 1153 GGCTATGAATTGATGCGTCCAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCCCTC 1212
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 1213 AACATTCAAGTGGCATCCTGCACAAAGGGAATGGGAGAACCCAGTTTGTGTCTCAGCAG 1272
QY 101 Pro 101
Db 1273 CCT 1275
RESULT 4
AY405422 5087 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405422
VERSION AY405422.1 GI:39761396
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 5087)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5087)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source Location/Qualifiers
1..5087
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>5087
/locus_tag="HCM2218"
gene
ORIGIN
Alignment Scores:
Pred. No.: 1.32e-50 Length: 5087
Score: 454.00 Matches: 79
Percent Similarity: 91.09% Conservative: 13
Best Local Similarity: 78.22% Mismatches: 9
Query Match: 82.70% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY405422 (1-5087)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 421 AGACTCTCCAGAAAGTGGTTTCCTGCCTCGCCAACTTGGCCCTACACGTTTCATCTGGGAT 480
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 481 AAGACCGACGCATATATCAGAAAGTCTACGGCTTGTTCAGAGGAGTGTGTCCGTCCGA 540
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 541 TACGAGTACGAGTCGTGCTTGGACCTGACTCTCTCTGGGAAAAGAGGACTGCCGTTTGGCA 600
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 601 GGCTATGAGTTGGATGCTTCGAACATGGGCGGCTGGACGTTGGACAAGCACCATGTACTG 660
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 661 GACGTTCAGAACGGTATACTATACAAAGGAAATGGAGAAATCAGTTTCATCTCTCAGCAG 720
QY 101 Pro 101
Db 721 CCT 723
RESULT 5
AY405421 5069 bp DNA linear GSS 12-DEC-2003
LOCUS Pan troglodytes HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405421
VERSION AY405421.1 GI:39761395
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 5069)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5069)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..5069
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>5069
/locus_tag="HCM2218"
gene
ORIGIN
Alignment Scores:
Pred. No.: 1.8e-50 Length: 5069
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY405421 (1-5069)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 422 AGACTCTTCCAAAGTGGTTTCCTGCCTCACCAAACCTTGGCCTATACTTTCATATGGGAT 481
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 482 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTGTCTGAAGCTGTTGTGTCAGTTGGA 541
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 542 TATGAGTATGAGTCGTGTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 601
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 602 GGCTATGAATTGGATGCGTCCAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 661
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 662 GATGTACAGAACGGTATACTGTACAAGGGAAACGGGGAAACCAGTTTCATCTCCCGCAG 721
QY 101 Pro 101
Db 722 CCT 724
RESULT 6
AY405420 5094 bp DNA linear GSS 12-DEC-2003
LOCUS Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
AY405420
ACCESSION AY405420.1 GI:39761394
VERSION
KEYWORDS GSS.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5094)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 5094)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..5094
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>5094
/locus_tag="HCM2218"
gene
ORIGIN
Alignment Scores:
Pred. No.: 1.81e-50 Length: 5094
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 29 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AY405420 (1-5094)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 422 AGACTCTTCCAAAGTGGTTTCCTGCCTCACCAAACCTTGGCCTATACTTTCATATGGGAT 481
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 482 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTCAGTTGGA 541
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 542 TATGAGTATGAGTCGTGTTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 601
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 602 GGCTATGAATTGGATGCGTCCAACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 661
QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 662 GATGTACAGAACGGTATACTGTACAAGGGAAACGGGGAAACCAGTTTCATCTCCCGCAG 721
QY 101 Pro 101
Db 722 CCT 724
RESULT 7
CD578516 796 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-GI0-cgg-e-23-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:6852960 5', mRNA sequence.
CD578516
ACCESSION CD578516.1 GI:31742907
VERSION

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 796)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
Seq primer: pYX-5.
FEATURES
source Location/Qualifiers
1..796
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clones="IMAGE:6852960"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_G10"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according Bonaïdo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 2,39e-47 Length: 796
Score: 422.00 Matches: 78
Percent Similarity: 89.11% Conservative: 12
Best Local Similarity: 77.23% Mismatches: 11
Query Match: 76.87% Indels: 1
DB: 14 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x CD578516 (1-796)
QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
464 AGACTCTCCAGAGTGGTTTCCTGCCTCGCCAAACTTGGCCTACACGTTTCATCTGGAT 523
Db
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
524 AAGACGGACGCATATAATCAGAAAGTCTACGGTTGTTCAGAGGCAGTTGTCTCGTCGGA 583
Db
QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
584 TACGAGTACGAGTCTGCTTGGACCTGACTCTCTGGGAAAGAGGACTGCGGTTTTCGA 643
Db
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
644 GGCTATGAGTTGGATGCTTCGAACATGGGGCGGTGGACGTTGGACAAGCA-CATGTACTG 702
Db

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 703 GACGTTCAAGACGGTACTATACAAAGGAAATGGAGAAATCAGTTCACTCTCAGCAG 762
QY 101 Pro 101
|||
Db 763 CCT 765
RESULT 8
AL710250
LOCUS DKFZp686F1868_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DEFINITION DKFZp686F1868 5', mRNA sequence.
ACCESSION AL710250
VERSION AL710250.1 GI:19693605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 750)
AUTHORS Ottenwaelder,B., Obermaier,B., Mewes,W., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE EST (Ottenwaelder,B., Obermaier,B., Mewes,H.W., Weil,B. and Wiemann,S.)
JOURNAL Unpublished (2001)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. No sl sequence available.
This clone (DKFZp686F1868) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source Location/Qualifiers
1..750
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686F1868"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB; cDNA-collection"
ORIGIN
Alignment Scores:
Pred. No.: 9.17e-45 Length: 750
Score: 403.00 Matches: 72
Percent Similarity: 87.13% Conservative: 16
Best Local Similarity: 71.29% Mismatches: 12
Query Match: 73.41% Indels: 1
DB: 9 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AL710250 (1-750)
QY 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 404 CTCCTCCAGAAGTCATTCACAGGCTTCTCCCAACCTGGCCTACACCTCATCTGGGACAAG 463
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 464 ACAGATCGGTATGGCCAAAGGGTGTATGGACTCTCAGATGCTGTGTCTCGGGTTT 523
QY 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 524 GAATATGAGACCTGTCCCAGTCTAATTCCTGGGAGAAAAGGACAGCCCTCCTTCAGGGA 583

QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTyrPsrLeuAspLysHisAlaLeuAsn 81
Db 584 TTCGAGCTGGACCCCTCCAACCTCGTGGTCCCTAGACAAACACCATCCTCAAT 643
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsn-GlyGluAsnGlnPheValSerGlnGlnPr 101
Db 644 GTTAAAGTGGAAATCCTACACAAAGGCACTGGGGAAACCAGTTCTCTGACCCAGCAGCC 703
QY 101 o 101
Db 704 T 704
RESULT 9
AL919858
LOCUS AL919858 PJR-Z1+Z2 Danio rerio cDNA clone 069-H12-2, mRNA sequence.
ACCESSION AL919858
VERSION AL919858.1 GI:23185156
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 539)
AUTHORS Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
Wang,W., Wen,Z. and Peng,J.
TITLE 15000 unique zebrafish EST clusters and their future use in
microarray for profiling gene expression patterns during
embryogenesis
JOURNAL Genome Res. 13 (3), 455-456 (2003)
MEDLINE 22505427
PUBMED 12618376
COMMENT Contact: Peng J
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengjr@imcb.a-star.edu.sg
Clone requests: pengjr@imcb.a-star.edu.sg.
FEATURES
source
Location/Qualifiers
1. .539
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="069-H12-2"
/tissue_type="whole embryo or fish"
/dev_stage="mixed stages"
/clone_lib="PJR-Z1+Z2"
ORIGIN
Alignment Scores:
Pred. No.: 2.58e-43 Length: 539
Score: 391.00 Matches: 68
Percent Similarity: 83.84% Conservative: 15
Best Local Similarity: 68.69% Mismatches: 16
Query Match: 71.22% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_1100_1200 (1-101) x AL919858 (1-539)
QY 2 LeupheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
Db 35 CTCCTCCAGAAATGGTTCCATGCCTCTCCTAACCTGGCATATACCTATATCTGGGACAAG 94
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
Db 95 ACTGATGCCTACAGGCAGAGGGTCCACGGCTTGACCGAAGCATTTGGTTTCTGGGGGTAC 154
QY 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
Db 155 GAGTATGAGACGTGTCCCGGTCAGCTCCTCTGTGGGAGAAAGGACAGCTGTGTGTGCAGGGA 214

QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTyrPsrLeuAspLysHisAlaLeuAsn 81
Db 215 TATGAGCTCAACCCCTCAAACCTTGGGCAGCTGGTCTCTGGACAAACACCATGCGGAAC 274
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 275 GTCCGCAGCGGAATCCTTCATAAGGGCAGCGGTGAAACATCTTTCTCTCTCAGCAA 331
RESULT 10
CA752216
LOCUS UI-M-FO0-cdn-g-07-0-UI.r1 NIH_BMAP_F00 Mus musculus cDNA clone
DEFINITION IMAGE:6830720 5', mRNA sequence.
ACCESSION CA752216
VERSION CA752216.1 GI:25582370
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 725)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
FEATURES
source
Location/Qualifiers
1. .725
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6830720"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_F00"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is TGAGAGAGCC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 5.44e-43 Length: 725
Score: 390.00 Matches: 68
Percent Similarity: 89.89% Conservative: 12
Best Local Similarity: 76.40% Mismatches: 9
Query Match: 71.04% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x CA752216 (1-725)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 459 AGACTCTCCAGAAGTGGTTCTCGCCTCGCCAAACTTGGCCTACACGTTTCATCTGGGAT 518

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 519 AAGACGGACGCATATAATCAGAAAAGTCTACGGCTTGTCAGAGGCAGTTGTGTCGTCGGA 578

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrValLeuGln 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 579 TACGAGTACGAGTCGTCGTTGGACCTGACTCTCTCGGAAAAGAGGACTGCCGTTTTCGAA 638

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 639 GGCATGAGTTGGATGCTTCGAACATGGGGCGCTGGACGTGGACAAGCACCATGTACTG 698

QY 81 AsnIleGlnSerGlyIleLeuHisLys 89
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 699 GACGTCAGAACGCTACTACTATACAAA 725

RESULT 11
CD803315 789 bp mRNA linear EST 15-JUL-2003
LOCUS UI-M-GVO-chu-m-02-0-UI.r1 NIH_BMAP_GVO Mus musculus cDNA clone
DEFINITION IMAGE:30545857 5', mRNA sequence.
ACCESSION CD803315
VERSION CD803315.1 GI:32462141
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 789)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. .789
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30545857"
/tissue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GVO"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated with
EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CGAACTGAAT. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 4.13e-42 Length: 789
Score: 384.00 Matches: 67
Percent Similarity: 91.57% Conservative: 9
Best Local Similarity: 80.72% Mismatches: 7
Query Match: 69.95% Indels: 0
DB: 14 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x CD803315 (1-789)

QY 19 TrpAspLysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSer 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 2 TGGGATAAGACGCGACGCATATAATCAGAAAGTCTACGGCTTGTCAGAGGCAGTTGTGTGCC 61

QY 39 ValGlyTyrGluTyrGluSerCysProAspLeuIleLeuTyrGluLysArgThrThrVal 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 62 GTCGATACGAGTACGAGTCGTCGTTGGACCTGACTCTCTGGGAAAAGAGGACTGCCGTT 121

QY 59 LeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisHis 78
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 122 TTGCAAGGCTATGAGTTGGATGCTTCGAACATGGGGCGCTGGACGTTGGACAAGCACCAT 181

QY 79 AlaLeuAsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSer 98
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 182 GTACTGGACGTTTCAGAACGGTACTACTATACAAAGGAATGGAGAAATCAGTTTCATCTCT 241

QY 99 GlnGlnPro 101
|||||:|||||
Db 242 CAGCAGCCT 250

RESULT 12
CB704746 348 bp mRNA linear EST 10-APR-2003
LOCUS AMGNNUC:SRPB2-00285-D10-A srpb2 (10220) Rattus norvegicus cDNA
DEFINITION clone srpb2-00285-d10 5', mRNA sequence.
ACCESSION CB704746
VERSION CB704746.1 GI:29761894
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 348)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00285 row: d column: 10.

FEATURES
source
1. .348
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="srpb2-00285-d10"
/tissue_type="prostate tissue"
/clone_lib="srpb2 (10220)"
/note="Vector: pSPORtl; Site_1: SalI; Site_2: NotI; rat
prostate normalized double selected poly(A+) mRNA size
fraction > 1 kb"

ORIGIN

Alignment Scores:
Pred. No.: 1.62e-41 Length: 348
Score: 376.00 Matches: 69
Percent Similarity: 87.50% Conservative: 15

SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 318)
AUTHORS
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
Kusakabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
Takahashi,F., Tominaga,N., Toya,T., Tsunoda,Y., Watahiki,A.,
Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.
TITLE
RIKEN Mouse ESTs (Konno,H., et al.)
JOURNAL
COMMENT
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Thermostabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for
further details.
FEATURES
source
Location/Qualifiers
1..318
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D630034A13"
/tissue type="kidney"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate
kidney"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGCGCACTCGAGTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATTCTCGAGTTAATAATTATCCCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

Pred. No.: 1.05e-29 Length: 318
Score: 290.00 Matches: 52
Percent Similarity: 73.63% Conservative: 15
Best Local Similarity: 57.14% Mismatches: 18
Query Match: 52.82% Indels: 6
DB: 10 Gaps: 1
US-10-029-020-14_COPY_1100_1200 (1-101) x BB501284 (1-318)
QY 4 ArgLysTrp-----PheAlaAlaProAspLeuSerTyrTyrPhe 17
||||:|||||
Db 46 CGTCGTGGGAGACTCTTCCACAACTCTTCTGCCTCCCTAACTTGCCTACACCTTT 105
QY 18 IleTrpAspLysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheVal 37
||||:|||||
Db 106 TTCTGGGATAAGACGAACCCATAATTTTCAGAAACTTTACGGCTTTTCAGAGGCACCTTGTT 165
QY 38 SerValGlyTyrGluTyrGluSerCysProAspLeuLeuLeuTrpGluLysArgThrThr 57
||||:|||||
Db 166 TCCCTCGGATACGAGTACGATTCTCTTGGACCTGACTCTCTGGTAAAGAGGACTCCC 225
QY 58 ValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHis 77
||||:|||||
Db 226 TTTTTCGAAGGCTATGAGTTGGATGCTTCGAACATGGGCGCTGGACGTTGGACAAGCAC 285
QY 78 HisAlaLeuAsnIleGlnSerGlyIleLeuHis 88
||||:|||||
Db 286 CATTTACTGGACGTTTCAGAACGGTATACTATAC 318
RESULT 15
BJ016612/c
LOCUS
BJ016612 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA030F08 3',
DEFINITION
mRNA sequence.
ACCESSION
BJ016612
VERSION
BJ016612.1 GI:17386535
KEYWORDS
EST.
SOURCE
Oryzias latipes (Japanese medaka)
ORGANISM
Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE
1 (bases 1 to 505)
AUTHORS
Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE
Medaka EST Project in Takeda's lab
JOURNAL
Unpublished (2001)
COMMENT
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
Location/Qualifiers
source
1..505
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01SSA030F08"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/clone_lib="MF01SSA cDNA"
ORIGIN
Alignment Scores:
Pred. No.: 5.27e-29 Length: 505
Score: 287.00 Matches: 52
Percent Similarity: 88.24% Conservative: 8
Best Local Similarity: 76.47% Mismatches: 8
Query Match: 52.28% Indels: 0

```
DB:                12                Gaps:                0
US-10-029-020-14_COPY_1100_1200 (1-101) x BJ016612 (1-505)
QY  34  GluAlaPheValSerValGlyTyrGluTyrGluSerCysProAspLeuIleLeuTrpGlu 53
Db  503 GAGGCAGTTGTGTCGGTGGGGTTTGAATACGAGTCCTGCCTTGACCTCATCTTGTGGGAG 444
QY  54  LysArgThrThrValLeuGlnGlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSer 73
Db  443 AAGAGGACGGCCATCTTACAAGGCTACGAGATGGACGCGTCCAAACATGGGAGGATGGACG 384
QY  74  LeuAspLysHisHisAlaLeuAsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGlu 93
Db  383 TTGGACAAACATCACATCCTGGATGTGCAAGCGGTATCTTGTATAAGGCAACGGTGAA 324
QY  94  AsnGlnPheValSerGlnGlnPro 101
Db  323 AATATCTTTGTATCCCTGCTACCT 300
```

Search completed: August 14, 2004, 18:03:09
Job time : 1197.51 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 171.323 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_1100_1200
Perfect score: 549
Sequence: 1 RLFRKWFAAAPDLSSYFIWD.....IQSGILHKGNGENQFVSQQP 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues 6747726
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO spool/US10029020/runat_06082004_112215_29265/app_query_fasta_1.3519
-DB=N Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1 1 1868 @runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DE TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002s:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	549	100.0	8354	6 ABS52100	Abs52100 Human TEN
2	549	100.0	8645	6 ABS78652	Abs78652 Human cDN
3	538.5	98.1	8438	6 ABN85378	Abn85378 Human NOV
4	532	96.9	453	3 AAC74626	Aac74626 Human ORF
5	453	82.5	8473	6 ABQ82345	Abq82345 Human NOV
6	453	82.5	8487	6 ABQ82346	Abq82346 Human NOV
7	453	82.5	8645	6 ABQ82344	Abq82344 Human NOV
8	453	82.5	8675	6 ABQ82343	Abq82343 Human NOV

9	416	75.8	4245	7 AAL60066	Aal60066 Human Pc0
10	416	75.8	9058	7 ACC72051	Acc72051 BCU0205A
11	416	75.8	9695	7 ACC72052	Acc72052 BCU0205B
12	411	74.9	12879	6 ABK92230	Abk92230 Prostate
13	411	74.9	13202	4 AAK51828	Aak51828 Human pol
14	407	74.1	9729	5 AAS14089	Aas14089 Human FCT
15	407	74.1	9729	9 ADB32028	Adb32028 Human FCT
16	407	74.1	9826	5 AAS14085	Aas14085 Human FCT
17	407	74.1	9826	9 ADB32023	Adb32023 Human FCT
18	341	62.1	466	7 ABT21688	Abt21688 Breast ca
19	248.5	45.3	10242	4 ABL29075	Abl29075 Drosophil
20	248.5	45.3	17131	4 ABL29074	Abl29074 Drosophil
21	234.5	42.7	1764	4 ABL04855	Abl04855 Drosophil
22	234.5	42.7	4925	4 ABL04854	Abl04854 Drosophil
23	73	13.3	2516	5 AAF89891	Aaf89891 Peroxisom
24	73	13.3	3008	6 ABK90574	Abk90574 Human PGC
25	73	13.3	3023	3 AAA07580	Aaa07580 Human PGC
26	73	13.3	3023	6 ABK90559	Abk90559 Human PGC
27	73	13.3	3023	8 ADA26907	Ada26907 Human PGC
28	73	13.3	6303	7 AAL55176	Aal55176 PGC-1 mut
29	73	13.3	6303	7 AAL55139	Aal55139 Wild-type
30	73	13.3	6306	3 AAC75134	Aac75134 Human ORF
31	71	12.9	2447	6 ABL90799	Abl90799 Human pol
32	71	12.9	2759	6 ABQ54635	Abq54635 Human ova
33	71	12.9	3653	6 ABS78730	Abs78730 DNA encod
34	70.5	12.8	699	6 ABN16991	Abn16991 Human ORF
35	70.5	12.8	866	6 ABK65370	Abk65370 Arabidops
36	70.5	12.8	866	9 ADD30950	Add30950 Plant yie
37	70.5	12.8	866	9 ADE37148	Ade37148 Plant yie
38	70.5	12.8	1169	3 AAC47819	Aac47819 Arabidops
39	70.5	12.8	4215	7 ABX34555	Abx34555 Human mdd
40	70	12.8	375	4 AAI85693	Aai85693 Human pol
41	70	12.8	816	2 AAT67351	Aat67351 H. pylori
42	70	12.8	855	2 AAT68114	Aat68114 H. pylori
43	70	12.8	984	4 AAS53729	Aas53729 Helicobac
44	70	12.8	984	4 AAS53891	Aas53891 Helicobac
45	70	12.8	2064	4 ABL19009	Ab119009 Drosophil

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.

XX AC ABS52100;

XX DT 05-NOV-2002 (first entry)

XX XX Human TEN-M4-like gene.

XX KW Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disorder; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related C1q tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophilin;
KW type 1a membrane-sushi domain containing; SNP; gene; ds;
KW single nucleotide polymorphism.

XX OS Homo sapiens.

XX XX Key Location/Qualifiers
FT variation replace(117,G)
FT /*tag= a


```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
PN WO200257453-A2.
XX
XX 25-JUL-2002.
XX
XX 19-DEC-2001; 2001WO-US050331.
XX
PR 19-DEC-2000; 2000US-0256704P.
PR 20-DEC-2000; 2000US-0257314P.
PR 02-MAY-2001; 2001US-0288153P.
PR 29-MAY-2001; 2001US-0294075P.
PR 24-JUL-2001; 2001US-0307506P.
PR 10-AUG-2001; 2001US-0311590P.
PR 10-AUG-2001; 2001US-0311613P.
PR 29-AUG-2001; 2001US-0315617P.
PR 14-SEP-2001; 2001US-0322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
DR WPI; 2002-590744/63.
DR P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
PT cancer.
XX
XX Claim 8; Page 50-52; 318pp; English.
XX
CC The present invention relates to new NOVX polypeptides. The invention is
CC useful for treating or preventing a NOVX-associated disorder such as
CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
CC signal processing and metabolic pathway modulation in a subject,
CC preferably human. The invention is also useful for treating metabolic
CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
CC disorders, haematopoietic disorders and various cancers. The molecules of
CC the invention are also useful for treating or preventing cirrhosis,
CC pancreatitis, learning and memory defects, infertility, congenital heart
CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
CC respiratory disease, gastro-intestinal diseases, reproductive, health,
CC neurological diseases, bone marrow transplantation, endocrine diseases,
CC allergy and inflammation, nephrological disorders, urinary system
CC disorders, neuropsychiatric disorders and age-related disorders. The
CC present nucleic acid sequence represents a NOVX gene. This sequence
CC encodes a NOVX protein of the invention
XX
SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.1e-66 Length: 8354
Score: 549.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ABS52100 (1-8354)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3332 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGCTCTATTATTCATTGGGAC 3391
```

```
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3392 AAGACAGACGCTCTACAACCAAGAGCGTGTGGGCTTTTCAGAGCCTTTGTTCCGTGGGT 3451

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 3452 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGTGCTGCAG 3511

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80
Db 3512 GGCTATGAAATTGACGCGTCCAGCTTGAGGATGGAGCCTAGACAAACATCATGCCCTC 3571

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db 3572 AACATTCAAAGTGGTATCCTGTCACAAAGGGAATGGGGAGACCAGTTTGTGTCTCAGCAG 3631

QY 101 Pro 101
Db 3632 CCT 3634

RESULT 2
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
AC ABS78652;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
KW polycythaemia vera; primary thrombocytopaenia; developmental disorder;
KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN WO200272830-A2.
XX
PD 19-SEP-2002.
XX
PF 08-FEB-2002; 2002WO-US003715.
XX
PR 09-FEB-2001; 2001US-0268111P.
PR 23-FEB-2001; 2001US-0271175P.
PR 08-MAR-2001; 2001US-0274503P.
PR 09-MAR-2001; 2001US-0274552P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
PI Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
DR WPI; 2002-723356/78.
DR P-PSDB; ABG97359.
XX
PT New human proteins associated with cell growth, differentiation and
PT death, useful for diagnosing, treating or preventing autoimmune or
PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
PT atherosclerosis or hepatitis.
XX
PS Claim 5; Page 175-178; 181pp; English.
XX
```

CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 5.35e-66 Length: 8645
Score: 549.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ABS78652 (1-8645)

Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db |||||||
3402 CGCCTCTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGCTATTATTTCATTGGGAC 3461
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db |||||||
3462 AAGACAGACGCTCTACAACCAGAGGTGTTTGGGCTTTCAGAGCCTTGTTCCTGGGT 3521
Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db |||||||
3522 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGTGTGTCAG 3581
Qy 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db |||||||
3582 GGCTATGAAATTGATGCGTCCAGCTTGGAGGATGGAGCCTAGACAAACATCATGCCCTC 3641
Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db |||||||
3642 AACATTCAAAGTGGCATCCTGCACAAGGGAATGGGGAGAACCATGTTGTGTCTCAGCAG 3701
Qy 101 Pro 101
Db |||||
3702 CCT 3704

RESULT 3
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
XX
AC ABN85378;
XX
DT 21-OCT-2002 (first entry)

XX Human NOV1, TEN-M4 like protein, coding sequence.
DE
XX
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IgA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX
OS Homo sapiens.
XX
FH
FT 4. .8395 Location/Qualifiers
FT /*tag= a
FT /trans_except= (pos: 1138. .1147,aa:Met)
FT /product= "NOV1 protein"
XX
WO200255704-A2.
XX
PD 18-JUL-2002.
XX
PF 09-JAN-2002; 2002WO-US0000554.
XX
PR 09-JAN-2001; 2001US-0260417P.
PR 10-JAN-2001; 2001US-0260831P.
PR 28-FEB-2001; 2001US-0272338P.
PR 09-MAR-2001; 2001US-0274876P.
PR 18-APR-2001; 2001US-0284704P.
XX
XX (CURA-) CURAGEN CORP.
PA Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
XX Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
PI Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
PI Gunther E, Stone DJ;
XX
DR WPI; 2002-590674/63.
DR P-PSDB; ABB98401.
XX
PT NOVX polypeptides and encoding polynucleotides, useful for preventing or
PT treating NOVX-associated disorders e.g. cancer, inflammation, or
PT Alzheimer's disease, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 9; Page 8-9; 358pp; English.
XX
CC The present sequence is a coding sequence for a NOV protein. The NOV
CC proteins and coding sequences are useful for treating or preventing NOV-
CC associated disorders or in the manufacture of a medicament for treating
CC the disorders, such as cancer, heart disease, inflammation, autoimmune
CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
CC IgA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
CC and other wasting disorders associated with chronic diseases. NOV1 is a
CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
XX
SQ Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.57e-64 Length: 8438
Score: 538.50 Matches: 101
Percent Similarity: 99.02% Conservative: 0
Best Local Similarity: 99.02% Mismatches: 0
Query Match: 98.09% Indels: 1
DB: 6 Gaps: 1

US-10-029-020-14_COPY_1100_1200 (1-101) x ABN85378 (1-8438)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 3311 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGTCCTATTATTTCATTGGGAC 3370
QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 3371 AAGACAGACGCTCTACAACCAGAGGTGTTGGGCTTTCAGAAAGCCTTGTTCCTGGGT 3430
QY 41 TyrGluTyrGluSerCysProAspLeuLeuLeuTrpGluLysArgThrThrValLeuGln 60
Db 3431 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAACAAACAGTGTGTCAG 3490
QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrPheValSerLysHisAlaLeu 80
Db 3491 GGCTATGAAATTGACGCGTCCAAAGCTTGAGGATGGAGCCTAGACAAACATCATGCCCTC 3550
QY 81 AsnIleGlnSer---GlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGln 99
Db 3551 AACATTCAAAGTGGTGGCATCCTGCACAAAGGAATGGGGAGAACCAAGTTTGTGTCTCAG 3610
QY 100 GlnPro 101
Db 3611 CAGCCT 3616

RESULT 4
AAC74626
ID AAC74626 standard; cDNA; 453 BP.
XX
AC AAC74626;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human ORFX ORF181 polynucleotide sequence SEQ ID NO:361.
XX
KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
KW vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KW cholesterol ester storage; systemic lupus erythematosus; infection;
KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
KW thrombosis; contraceptive; ss.

Homo sapiens.
OS
XX
PN WO200058473-A2.
XX
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US008621.
XX
PR 31-MAR-1999; 99US-0127607P.
PR 02-APR-1999; 99US-0127636P.
PR 05-APR-1999; 99US-0127728P.
PR 30-MAR-2000; 2000US-00540763.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2000-602362/57.
DR P-PSDB; AAB40417.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease.

XX
PS
XX
CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;
CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
CC antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The
CC sequences can be used for determining the presence of or predisposition
CC to, or preventing or treating pathological conditions associated with an
CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
CC proteins in gene therapy vectors. The proteins and nucleic acids may be
CC used to treat cancers, proliferative disorders, neurodegenerative
CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
CC storage, systemic lupus erythematosus, severe combined immunodeficiency
CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
XX
SQ Sequence 453 BP; 115 A; 117 C; 116 G; 104 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 2.2e-65 Length: 453
Score: 532.00 Matches: 98
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.90% Indels: 0
DB: 3 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AAC74626 (1-453)
Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db 160 CGCCTCTTCAGGAAGTGGTTCGCTGCAGCCCCAGACCTGCTCTATTATTTCATTGGGAC 219
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db 220 AAGACAGACGCTCTACAACCAGAGGTGTTGGGCTTTCAGAAAGCCTTGTTCCTGGGT 279
Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db 280 TATGAATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAGAACAAACAGTGTGTCAG 339
Qy 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrPheValSerLysHisAlaLeu 80
Db 340 GGCTATGAAATTGACGCGTCCAGCTTGAGGATGGAGCTAGACAAACATCATGCCCTC 399
Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSer 98
Db 400 AACATTCAAAGTGGCATCCTGCACAAAGGAATGGNGAGAACCCAGTTTGTGTCT 453

RESULT 5
ABQ82345
ID ABQ82345 standard; cDNA; 8473 BP.
XX
AC ABQ82345;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15c encoding cDNA SEQ ID NO:39.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;

KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 258..8144
FT /*tag= a
FT /product= "NOV15c"
XX

PN WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX

PA (CURA-) CURAGEN CORP.

XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;

DR WPI; 2002-732706/79.
DR P-PSDB; ABP53588.

XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.

XX Claim 8; Page 119-121; 444pp; English.

XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytotstatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a

CC biological sample. The present sequence encodes human NOV15c, which is
CC located on chromosome 4
XX
SQ Sequence 8473 BP; 2301 A; 2017 C; 2145 G; 2006 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 1.83e-52 Length: 8473
Score: 453.00 Matches: 78
Percent Similarity: 91.09% Conservative: 14
Best Local Similarity: 77.23% Mismatches: 9
Query Match: 82.51% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ABQ82345 (1-8473)

Qy 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20
Db |||||:::||||| ||:::||||:||||:||||:||||:||||:||||:||||
3150 AGACTCTTCCAAAAGTGGTTCTCTGCCTCACCAAACTGGCCTATACTTTTCATATGGGAT 3209
Qy 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40
Db |||||:::||||| |||||:::||||:||||:||||:||||:||||:||||
3210 AAAACAGATGCATATATCAGAAAAGTCTATGGTCTATCTGAAGCTGTTGTGTCAGTTGGA 3269
Qy 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60
Db |||||:::||||| |||||:::||||:||||:||||:||||:||||:||||
3270 TATGAGTATGAGTCGTGTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATTCTGCAG 3329
Qy 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeu 80
Db |||||:::||||| |||||:::||||:||||:||||:||||:||||:||||
3330 GGCATGAATTGGATGCGTCCAACATGGTGGCTGGCATTAGATAAACATCACGTGCTG 3389
Qy 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100
Db :::::||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
3390 GATGTACAGAACGGTATACTGTACAAGGGAAACGGGGAAAACCAAGTTTCATCTCCAGCAG 3449

Qy 101 Pro 101

Db 3450 CCT 3452

RESULT 6

ABQ82346

ID ABQ82346 standard; cDNA; 8487 BP.

XX ABQ82346;

XX 17-DEC-2002 (first entry)

DE Human NOV15d encoding cDNA SEQ ID NO:41.

XX
KW Human; NOVX; cytotstatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

OS

XX Key Location/Qualifiers

FH CDS 299..8140

FT /*tag= a

FT /product= "NOV15d"

XX WO200262999-A2.

PN 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

PF.

PA (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

PI Gunther E, Smithson G, Millet I, Macdougall JR;

XX

DR WPI; 2002-732706/79.

DR P-PSDB; ABP53587.

XX

PT New NOVX polypeptides and polynucleotides useful for treating NOVX-

PT associated disorders, such as cancers, neurological disorders, disorders

PT of vesicular transport, gastrointestinal disorders, and autoimmune

PT diseases.

XX

PS Claim 8; Page 114-117; 444pp; English.

XX

CC The present invention describes novel human proteins designated NOVX,

CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,

CC cytotstatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,

CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,

CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,

CC protozoacide and antihelminthic activities, and can be used in gene

CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the

CC manufacture of a medicament for treating a syndrome associated with a

CC human disease selected from NOVX-associated disorder, such as cancers

CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,

CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,

CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's

CC disease), disorders of vesicular transport (e.g. cystic fibrosis,

CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal

CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),

CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic

CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic

CC and protozoal infections. The NOVX proteins can be used as immunogens to

CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be

CC used in chromosome mapping, identifying individuals from minute

CC biological samples (tissue typing), and in forensic identification of a

CC biological sample. The present sequence encodes human NOV15b, which is

CC located on chromosome 4

XX

SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.:	1-88e-52	Length:	8645
Score:	453.00	Matches:	78
Percent Similarity:	91.09%	Conservative:	14
Best Local Similarity:	77.23%	Mismatches:	9
Query Match:	82.51%	Indels:	0
DB:	6	Gaps:	0

US-10-029-020-14_COPY_1100_1200 (1-101) x ABQ82344 (1-8645)

QY 1 ArgLeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAsp 20

DB 3322 AGACTCTTCCAAAGTGTTTCTGCTCACCACCAACTTGGCCTATACTTTCATATGGGAT 3381

QY 21 LysThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGly 40

DB 3382 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTGTGAGTTGGA 3441

QY 41 TyrGluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGln 60

DB 3442 TATGATGATGAGTCGTGTTTGGACCTGACTCTGTGGGAAAAGAGACTGCCATTCTGCAG 3501

QY 61 GlyTyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeu 80

DB 3502 GGCTATGAATTGGATGCGTCCACATGGGTGGCTGGACATTAGATAAACATCACGTGCTG 3561

QY 81 AsnIleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGln 100

DB 3562 GATGTACAGAACGGTATACTGTACAGGGAACCGGGGAAAACAGATTTCATCTCCCGACGAG 3621

QY 101 pro 101

DB 3622 CCT 3624

RESULT 8

ABQ82343

ID ABQ82343 standard; cDNA; 8675 BP.

XX

AC ABQ82343;

XX

DT 17-DEC-2002 (first entry)

XX

DE Human NOV15a encoding cDNA SEQ ID NO:35.

XX

KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;

KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;

KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;

KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;

KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;

KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;

KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;

KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;

KW ulcerative colitis; gastric disorder; duodenal disorder; infection;

KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;

KW rheumatoid arthritis; gene; chromosome 4; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 151..8328

FT /*tag= a

FT /product= "NOV15a"

FT /transl_except= (pos:1249..1251,aa:Ser)

XX

PN WO200262999-A2.

XX

PD 15-AUG-2002.

XX

PF 31-DEC-2001; 2001WO-US049976.

XX

PR 29-DEC-2000; 2000US-0258928P.

PR 02-JAN-2001; 2001US-0259415P.

PR 04-JAN-2001; 2001US-0259785P.

PR 20-FEB-2001; 2001US-0269814P.

PR 09-MAR-2001; 2001US-0279863P.

PR 29-MAR-2001; 2001US-0279832P.

PR 29-MAR-2001; 2001US-0279833P.

PR 13-APR-2001; 2001US-0283889P.

PR 18-APR-2001; 2001US-0284447P.

PR 25-APR-2001; 2001US-0286683P.

PR 29-MAY-2001; 2001US-0294080P.

PR 16-AUG-2001; 2001US-0312915P.

PR 17-AUG-2001; 2001US-0313325P.

PR 17-SEP-2001; 2001US-0322699P.

PR 26-NOV-2001; 2001US-0333350P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;

PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;

PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;

PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;

PI Gunther E, Smithson G, Millet I, Macdougall JR;

XX

DR WPI; 2002-732706/79.

DR P-PSDB; ABP53586.

XX

PT New NOVX polypeptides and polynucleotides useful for treating NOVX-

PT associated disorders, such as cancers, neurological disorders, disorders

PT of vesicular transport, gastrointestinal disorders, and autoimmune

PT diseases.

XX

QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
:::|||||
Db 2158 GTTAAAGTGGAAATCTACACAAAGGCATGGGAAACCAAGTTCTGACCCAGAGCCT 2217

RESULT 10
ACC72051
ID ACC72051 standard; DNA; 9058 BP.
XX
AC ACC72051;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205A gene #SEQ ID 79.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
PN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58317.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.34e-47 Length: 9058
Score: 416.00 Matches: 72
Percent Similarity: 88.00% Conservative: 16
Best Local Similarity: 72.00% Mismatches: 12
Query Match: 75.77% Indels: 0
DB: 7 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ACC72051 (1-9058)

QY 2 LeupheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
|||||:::|||||
Db 2699 CTCCTCCAGAAGTCATTCAGGCTTCTCCCAACCTGGCCTACACCTTCATCTGGACAAAG 2758
|||:::|||||
QY 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
|||||:::|||||

Db 2759 ACAGATCGGTATGGCCAAAGGTTGATGGAATCTCAGATGCTGTGTGTGTCGGGTTT 2818
|||||:::|||||
QY 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
|||||:::|||||
Db 2819 GAATATGAGACCTGTCCCAGTCTAATCTCTGGGAGAAAAGGACAGCCCTCTTCAGGGA 2878
|||||:::|||||
QY 62 TyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisAlaLeuAsn 81
:::|||||
Db 2879 TTCGAGCTGGACCCCTCCAACCTCGGTGGTGGTCCCTAGACAAACACCATCCTCAAT 2938
|||||:::|||||
QY 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
:::|||||
Db 2939 GTTAAAGTGGAAATCTACACAAAGGCACTGGGGAACCAAGTTCTGACCCAGAGCCT 2998
|||||:::|||||

RESULT 11
ACC72052
ID ACC72052 standard; DNA; 9695 BP.
XX
AC ACC72052;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205B gene #SEQ ID 81.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
PN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58318.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9695 BP; 2567 A; 2582 C; 2434 G; 2112 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.67e-47 Length: 9695
Score: 416.00 Matches: 72
Percent Similarity: 88.00% Conservative: 16

CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Groenouw type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC the FCTR3a homologue FCTR3f
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.86e-46 Length: 9729
Score: 407.00 Matches: 71
Percent Similarity: 87.00% Conservative: 16
Best Local Similarity: 71.00% Mismatches: 13
Query Match: 74.13% Indels: 0
DB: 5 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x AAS14089 (1-9729)

Qy 2 LeuPheArgLysTrpPheAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
Db 3384 CTCTTCCAGAGTCATTCCAGGCTTCTCCCAACCTGGCCTCCACCTTCATCTGGACAAG 3443
Qy 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
Db 3444 ACAGATGCGTATGGCCAAAGGTGTATGGACTCTCAGATGCTGTGTCTGTCGGGTTT 3503
Qy 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
Db 3504 GAATATGAGACCTGTCCAGTCTAATTCTCTGGGAGAAAGGACAGCCCTCCTTCAGGGA 3563
Qy 62 TyrGluIleAspAlaSerLysLeuGlyGlyTrpSerLeuAspLysHisAlaLeuAsn 81
Db 3564 TTCGAGCTGGACCCCTCCAACTCGTGGTGGTCCCTAGACAAACACCACATCTCAAT 3623
Qy 82 IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro 101
Db 3624 GTTAAAGTGAATCCTACACAAAGGCACTGGGAAACCAGTTCTTGACCCAGCCT 3683

RESULT 15
ADB32028
ID ADB32028 standard; cDNA; 9729 BP.

XX ADB32028;
XX
DT 04-DEC-2003 (first entry)
XX Human FCTR3f cDNA.
XX Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.

XX Homo sapiens.
OS
XX
PN US2003087816-A1.

XX 08-MAY-2003.
PD
XX
PF 05-MAR-2001; 2001US-00800198.
XX
PR 03-MAR-2000; 2000US-0186592P.
XX
PA (VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.
PA (RAST/) RASTELLI L.
XX
PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2003-625633/59.
DR P-PSDB; ADB32029.
XX

PT New FCTRX polypeptide and encoding polynucleotide, useful for preventing
PT or treating FCTRX-related disorders, such as cancer, autoimmune,
PT neurodegenerative, gastrointestinal, reproductive and inflammatory
PT diseases.
XX
PS Claim 9; Page 34-37; 155pp; English.

XX The invention relates to FCTRX polypeptides and the polynucleotides
CC encoding them. The sequences of the invention are useful for the
CC manufacture of a medicament for diagnosing and treating disorders
CC associated with the FCTRX polypeptide, such as colorectal cancer,
CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
CC cell and granular cell carcinomas, neurological disorders,
CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
CC conditions, immunological disorders, allergy and infection, asthma, lung
CC diseases, male and female reproductive disorders, deafness, glycoprotein
CC deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
CC C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
CC infection, spinocerebellar ataxia, plasmodium falciparum infection,
CC Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
CC represents cDNA encoding an FCTRX polypeptide of the invention.

XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6.86e-46 Length: 9729
Score: 407.00 Matches: 71
Percent Similarity: 87.00% Conservative: 16
Best Local Similarity: 71.00% Mismatches: 13
Query Match: 74.13% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_1100_1200 (1-101) x ADB32028 (1-9729)

Qy 2 LeuPheArgLysTrpPheAlaAlaAlaProAspLeuSerTyrTyrPheIleTrpAspLys 21
Db 3384 CTCTTCCAGAGTCATTCCAGGCTTCTCCCAACCTGGCCTCCACCTTCATCTGGACAAG 3443
Qy 22 ThrAspValTyrAsnGlnLysValPheGlyLeuSerGluAlaPheValSerValGlyTyr 41
Db 3444 ACAGATGCGTATGGCCAAAGGTGTATGGACTCTCAGATGCTGTGTCTGTCGGGTTT 3503
Qy 42 GluTyrGluSerCysProAspLeuIleLeuTrpGluLysArgThrThrValLeuGlnGly 61
Db 3504 GAATATGAGACCTGTCCAGTCTAATTCTCTGGGAGAAAGGACAGCCCTCCTTCAGGGA 3563

QY	62	TyrGluIleAspAlaSerLysLeuGlyGlyTyrSerLeuAspLysHisHisAlaLeuAsn	81
Db	3564	TTCGAGCTGGACCCCTCCAAACCTCGGTGGCTGGTCCCTAGACAAACACCATCCTCAAT	3623
QY	82	IleGlnSerGlyIleLeuHisLysGlyAsnGlyGluAsnGlnPheValSerGlnGlnPro	101
Db	3624	GTTAAAAGTGGAAATCCTACACAAAGGCACTGGGGAAAAACCAAGTTCCTGACCCAGCAGCCT	3683

Search completed: August 14, 2004, 02:29:50
Job time : 197.323 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 202.425 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WMGACDQACHPRCAEHGT.....DTSMETACGDSKNDGDGLV 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues
Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020 @CGN 1 1 2156 @runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1
US-10-383-201-49
; Sequence 49, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

1	617	100.0	1399	13	US-10-383-201-49	Sequence 49, Appl
2	617	100.0	1399	13	US-10-383-201-59	Sequence 59, Appl
3	617	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
4	617	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
5	595	96.4	8355	13	US-10-383-201-55	Sequence 55, Appl
6	579	93.8	8438	13	US-10-042-865-1	Sequence 1, Appli
7	555.5	90.0	1371	13	US-10-383-201-51	Sequence 51, Appl
8	515	83.5	8473	17	US-10-038-854-39	Sequence 39, Appl
9	515	83.5	8645	17	US-10-038-854-37	Sequence 37, Appl
10	515	83.5	8675	17	US-10-038-854-35	Sequence 35, Appl
11	509	82.5	829	13	US-10-383-201-47	Sequence 47, Appl
12	509	82.5	829	13	US-10-383-201-61	Sequence 61, Appl
13	509	82.5	1392	13	US-10-383-201-45	Sequence 45, Appl
14	509	82.5	1392	13	US-10-383-201-53	Sequence 53, Appl
15	509	82.5	1476	13	US-10-383-201-41	Sequence 41, Appl
16	478.5	77.6	8487	17	US-10-038-854-41	Sequence 41, Appl
17	464.5	75.3	12880	16	US-10-295-027-927	Sequence 927, App
18	454	73.6	9826	9	US-09-808-602-7	Sequence 7, Appli
19	454	73.6	9826	10	US-09-800-198-7	Sequence 7, Appli
20	452	73.3	8409	9	US-09-808-602-79	Sequence 79, Appl
21	452	73.3	8409	10	US-09-800-198-67	Sequence 67, Appl
22	448.5	72.7	8689	9	US-09-808-602-78	Sequence 78, Appl
23	448.5	72.7	8689	10	US-09-800-198-66	Sequence 66, Appl
24	448.5	72.7	8797	9	US-09-808-602-74	Sequence 74, Appl
25	448.5	72.7	8797	9	US-09-808-602-77	Sequence 77, Appl
26	448.5	72.7	8797	10	US-09-800-198-62	Sequence 62, Appl
27	448.5	72.7	8797	10	US-09-800-198-65	Sequence 65, Appl
28	448.5	72.7	9729	9	US-09-808-602-12	Sequence 12, Appl
29	448.5	72.7	9729	10	US-09-800-198-12	Sequence 12, Appl
30	447.5	72.5	802	13	US-10-383-201-57	Sequence 57, Appl
31	447.5	72.5	8575	13	US-10-072-012-143	Sequence 143, App
32	430.5	69.8	447	9	US-09-833-381-1072	Sequence 79, Appl
33	330.5	53.6	9058	16	US-10-144-194A-79	Sequence 81, Appl
34	330.5	53.6	9695	16	US-10-144-194A-81	Sequence 75, Appl
35	216	35.0	2496	9	US-09-808-602-75	Sequence 53, Appl
36	216	35.0	2496	10	US-09-800-198-63	Sequence 24765, A
C 37	199	32.3	260	15	US-10-029-386-24765	Sequence 11065, A
C 38	199	32.3	600	15	US-10-029-386-11065	Sequence 47, Appl
39	169	27.4	13268	15	US-10-007-926A-47	Sequence 75, Appl
40	169	27.4	13857	15	US-10-037-270-75	Sequence 75, Appl
41	169	27.4	13857	16	US-10-117-722-75	Sequence 1470, Ap
42	169	27.4	100267	13	US-10-240-425-1470	Sequence 10449, A
43	161	26.1	536	15	US-10-029-386-10449	Sequence 24152, A
44	161	26.1	1251	15	US-10-029-386-24152	Sequence 17, Appl
45	156	25.3	7560	10	US-09-921-406C-17	

```

; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 49
; LENGTH: 1399
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1399)
;
ITS-10-383-201-49

```

Alignment Scores:		
Pred. No.:	9.8e-64	1399
Score:	617.00	101
Percent Similarity:	100.00%	0
Best Local Similarity:	100.00%	0
Query Match:	100.00%	0
DB:	13	0
	Length:	
	Matches:	
	Conservative:	
	Mismatches:	
	Indels:	
	Gaps:	

Qy	1	TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr	20
Db	578	TGGATGGGGGCAGCCTGCGACCAAGCGGGCCTGCCACCCGCGCTGTGCCAGCATGGGACC	637
Qy	21	CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla	40
Db	638	TGCCGCGACGGCAAGTCGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT	697
Qy	41	HisTyrLeuAspArgValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg	60
Db	698	CACATATCTGGATAGGCTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGSCAACGGCAGA	757
Qy	61	CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly	80
Db	758	TGTACCTTAGACCTGAATGGTTGGCACTGCGCTCTGCCAGCTGGGCTGGAGAGGAGCTGGC	817
Qy	81	CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu	100
Db	818	TGTGACACTTCCATGGAGACTGCGCTGCGGTGACAGCAAGACAATGATGGAGATGGCCTG	877
Qy	101	Val 101	
Db	878	GTG 880	

```

; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 59

```

Alignment Scores:	
Pred. No.:	9.8e-64
Score:	617.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	13
Length:	1399
Matches:	101
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

QY	1	TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr	20
Db	578	TGGATGGGGGCAGCCTGCGACACAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACC	637
QY	21	CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla	40
Db	638	TGCCGCGACGGCAAGTGCAGAGTCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT	697
QY	41	HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg	60
Db	698	CACATATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGSCAACGGCAGA	757
QY	61	CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly	80
Db	758	TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC	817
QY	81	CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu	100
Db	818	TGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGACAAATGATGGAGATGGCCTG	877
QY	101	Val 101	
Db	878	GTG 880	

; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores:
Pred. No.: 7.93e-63 Length: 8354
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-43 (1-8354)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||||
Db 2282 TGGATGGGGCAGCCTGCGACCGGGCCTGCCACCCCGCGCTGTGCCGAGCATGGGACC 2341
|||
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||||
Db 2342 TGCCGCGACGGCAAGTCGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT 2401
|||
QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||||
Db 2402 CACTATCTGGATAGGTTAGTTAAAGAGGGTTGCCCTGGGTGTGCAATGGCAACGGCAGA 2461
|||
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||||
Db 2462 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC 2521
|||
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
|||||
Db 2522 TGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAGACAATGATGGAGATGGCCTG 2581
|||
QY 101 Val 101
|||
Db 2582 GTG 2584

RESULT 4
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Ganggilli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613

; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores:
Pred. No.: 7.93e-63 Length: 8354
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-029-020-13 (1-8354)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||||
Db 2282 TGGATGGGGCAGCCTGCGACCGGGCCTGCCACCCCGCGCTGTGCCGAGCATGGGACC 2341
|||
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||||
Db 2342 TGCCGCGACGGCAAGTCGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT 2401
|||
QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||||
Db 2402 CACTATCTGGATAGGTTAGTTAAAGAGGGTTGCCCTGGGTGTGCAATGGCAACGGCAGA 2461
|||
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||||
Db 2462 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC 2521
|||
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
|||||
Db 2522 TGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAGACAATGATGGAGATGGCCTG 2581
|||
QY 101 Val 101
|||
Db 2582 GTG 2584

RESULT 5
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26

; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55

Alignment Scores:
Pred. No.: 3.29e-60 Length: 8355
Score: 595.00 Matches: 101
Percent Similarity: 96.19% Conservative: 0
Best Local Similarity: 96.19% Mismatches: 0
Query Match: 96.43% Indels: 4
DB: 13 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-55 (1-8355)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2239 TGGATGGGGCAGCCTGCGACCGGGGCTGCCACCCGCGCTGTGCCGAGCATGGGACC 2298
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle--- 39
Db 2299 TGCCGCGACGGCAAGTGGAGTGGAGCCCTGGCTGGAATGGCGAACACTGCACCATCTCC 2358
QY 40 ---AlaHisTyrLeuAspArgValValLys-----GluGlyCysProGlyLeuCysAsn 56
Db 2359 CTAGCTCACTATCTGGATAGGGTAGTTAAACTTTTCAGAGGGTTGCCCTGGTTGTGCAAT 2418
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 2419 GGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGG 2478

QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAsp 96
Db 2479 AGAGGAGCTGGCTGTGACACTTCATGGAGACTGCTGCGGTGACAGCAAAAGACAATGAT 2538
QY 97 GlyAspGlyLeuVal 101
Db 2539 GGAGATGGCCTGGTG 2553

RESULT 6
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T

; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1

Alignment Scores:
Pred. No.: 2.67e-58 Length: 8438
Score: 579.00 Matches: 98
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 2
Query Match: 93.84% Indels: 4
DB: 13 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x US-10-042-865-1 (1-8438)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2249 TGGACGGGCCAGCCTGTAAATCAGAGAGCCTGCCACCCCGCTGTGCCGAGCACGGGACC 2308
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle--- 39
Db 2309 TGCCGCGACGGCAAGTGGAGTGGAGCCCTGGCTGGAATGGCGAACACTGCACCATCTCC 2368
QY 40 ---AlaHisTyrLeuAspArgValValLys-----GluGlyCysProGlyLeuCysAsn 56
Db 2369 CTAGCTCACTATCTGGATAGGGTAGTTAAACTTTTCAGAGGGTTGCCCTGGTTGTGCAAT 2428
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 2429 GGCAACGGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGG 2488
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAsp 96
Db 2489 AGAGGAGCTGGCTGTGACACTTCATGGAGACTGCTGCGGTGACAGCAAAAGACAATGAT 2548
QY 97 GlyAspGlyLeuVal 101
Db 2549 GGAGATGGCCTGGTG 2563

```
RESULT 7
US-10-383-201-51
; Sequence 51, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 51
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1369)
US-10-383-201-51

Alignment Scores:
Pred. No.:          2e-56          Length:      1371
Score:              555.50          Matches:      92
Percent Similarity: 91.09%          Conservative: 0
Best Local Similarity: 91.09%          Mismatches:  0
Query Match:        90.03%          Indels:       9
DB:                  13              Gaps:         1

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-51 (1-1371)

QY      1  TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
      |||
Db      578 TGGATGGGGGCAGCCTGCGACCAAGCGGGCCTGCCACCCCGCTGTGCCGAGCATGGGACC 637

QY      21  CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
      |||
Db      638 TGCCGCGACGGCAAGTGGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATC--- 694

QY      41  HisTyrLeuAspArgValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
      |||
Db      695 -----GAGGGTTGCCCTGGGTTGTGCAATGGCAATGGCAGA 730

QY      61  CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
      |||
Db      731 TGTACCTTAGACCTGAATGTTGGCACTGCGTCTGCCAGCTGGCTGGAGAGGAGCTGGC 790

QY      81  CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
      |||
Db      791 TGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAGACAATGATGGAGATGGCCTG 850

QY      101 Val 101
      |||
```

```
Db      851 GTG 853

RESULT 8
US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 8473
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-39

Alignment Scores:
Pred. No.:          1.11e-50          Length:      8473
Score:              515.00          Matches:      84
Percent Similarity: 85.05%          Conservative:  7
Best Local Similarity: 78.50%          Mismatches: 10
Query Match:        83.47%          Indels:       6
```



```
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 8675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-35
```

```
Alignment Scores:
Pred. No.: 1.14e-50 Length: 8675
Score: 515.00 Matches: 84
Percent Similarity: 85.05% Conservative: 7
Best Local Similarity: 78.50% Mismatches: 10
Query Match: 83.47% Indels: 6
DB: 17 Gaps: 1
```

US-10-029-020-14_COPY_750_850 (1-101) x US-10-038-854-35 (1-8675)

```
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2266 TGGACGGGCCCCAGCCTGTAATCAGAGAGCCTGCCACCCCGCTGTGCCGAGCAGCGGACC 2325
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2326 TGCAAGGATGGCAAGTGTAATGCAGCCAGGGCTGGAATGGAGAGCACTGCACATATCGCT 2385
QY 41 HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
Db 2386 CACTATTGGATAAGATAGTAAAGACAAGATAGGATATAAAGAGGGTTGTCTCGTCTG 2445
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 2446 TGCAACAGCAATGGAAGATGTACCTTGGACCAAAATGGCGGACATTGTGTGTGCCAGCCT 2505
QY 75 GlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
```

```
Db 2506 GGATGGAGAGGAGCAGGCTGTACGTAGCCATGGAGACTCTTTGCACAGATAGCAAGGAC 2565
QY 95 AsnAspGlyAspGlyLeuVal 101
Db 2566 AATGAAGGGGATGGACTCATT 2586
RESULT 11
US-10-383-201-47
; Sequence 47, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 47
; LENGTH: 829
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(829)
US-10-383-201-47
```

```
Alignment Scores:
Pred. No.: 3.8e-51 Length: 829
Score: 509.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.50% Indels: 0
DB: 13 Gaps: 0
```

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-47 (1-829)

```
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 578 TGGATGGGGGACGCTCGACCCAGCGGGCCTGCCACCCCGCTGTGCCGAGCATGGGACC 637
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 638 TGCCCGCAGCGCAAGTGCAGTGCAGCCCTGGCTGGGAATGGCGAACACTGCACCATCGCT 697
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 698 CACTATCTGATAGGCTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 757
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 758 TGTACCTTAGACCTGAATGGTTGGCACTGCGCTGCCAGCTGGGCTGGAGAGGAGCTGGC 817
```

QY 81 Cys 81
|||
Db 818 TGT 820

RESULT 12
US-10-383-201-61
; Sequence 61, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 61
; LENGTH: 829
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(820)
US-10-383-201-61

Alignment Scores:
Pred. No.: 3.8e-51 Length: 829
Score: 509.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.50% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-61 (1-829)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||
Db 578 TGGATGGGGGCAGCCTGCGACCAAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACC 637

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||
Db 638 TGCCGCGACGGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAACAACACTGCACCATCGCT 697

QY 41 HistyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||
Db 698 CACTATCTGGATAGGTTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 757

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||
Db 758 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC 817

QY 81 Cys 81
|||

Db 818 TGT 820
|||

RESULT 13
US-10-383-201-45
; Sequence 45, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 45
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1392)
US-10-383-201-45

Alignment Scores:
Pred. No.: 6.97e-51 Length: 1392
Score: 509.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.50% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-45 (1-1392)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||
Db 1144 TGGATGGGGGCAGCCTGCGACCAAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACC 1203

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||
Db 1204 TGCCGCGACGGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAACAACACTGCACCATCGCT 1263

QY 41 HistyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||
Db 1264 CACTATCTGGATAGGTTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 1323

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||
Db 1324 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC 1383

QY 81 Cys 81
|||

Db 1384 TGT 1386

RESULT 14
US-10-383-201-53
; Sequence 53, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 53
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1386)
US-10-383-201-53

Alignment Scores:
Pred. No.: 6.97e-51 Length: 1392
Score: 509.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.50% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-53 (1-1392)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 1144 TGGATGGGGCAGCCTGCGACCGGGCCTGCCACCGCGCTGTGCCGAGCATGGGACC 1203

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 1204 TGCCGCGACGGCAAGTGCAGTGCAGCCCTGGCTGGATGGCGAACACTGCACCATCGCT 1263

QY 41 HisTyrLeuAspArgValLysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 60
Db 1264 CACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGTTGTGCAATGGCAACGGCAGA 1323

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 1324 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC 1383

QY 81 Cys 81
Db 1384 TGT 1386

RESULT 15

US-10-383-201-41
; Sequence 41, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 41
; LENGTH: 1476
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1464)
US-10-383-201-41

Alignment Scores:
Pred. No.: 7.47e-51 Length: 1476
Score: 509.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 82.50% Indels: 0
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x US-10-383-201-41 (1-1476)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 1222 TGGATGGGGCAGCCTGCGACCGGGCCTGCCACCGCGCTGTGCCGAGCATGGGACC 1281

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 1282 TGCCGCGACGGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT 1341

QY 41 HisTyrLeuAspArgValLysGluGlyCysProGlyCysAsnGlyLeuCysAsnGlyArg 60
Db 1342 CACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 1401

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 1402 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC 1461

QY 81 Cys 81
Db 1462 TGT 1464

Search completed: August 14, 2004, 19:17:13
Job time : 216.425 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:16:21 ; Search time 29.8807 Seconds
(without alignments)
1875.793 Million cell updates/sec

Title: ,/ US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WMGAACDQACRCHPRCAEHGT.....DTSMETACGDSKNDGDGLV 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112217_29301/app_query.fasta_1.3519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Dits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn 1 1 258 @runat_06082004_112217_29301 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	430.5	69.8	447	US-09-833-381-1072	Sequence 1072, Ap
2	169.5	27.5	5561	US-08-400-159-1	Sequence 1, Appli
3	169.5	27.5	5561	US-08-611-729A-1	Sequence 1, Appli
4	169	27.4	13857	US-09-620-312D-75	Sequence 75, Appl
5	160	25.9	6049	US-08-793-273C-3	Sequence 3, Appli
6	160	25.9	6049	PCT-US95-11684-3	Sequence 3, Appli
7	155	25.1	4208	US-09-214-278-6	Sequence 6, Appli
8	155	25.1	4208	US-09-068-740A-10	Sequence 10, Appl
9	155	25.1	4208	US-09-855-722-6	Sequence 6, Appli
10	155	25.1	5458	US-09-199-865-2	Sequence 2, Appli
11	155	25.1	5590	US-08-882-046-1	Sequence 1, Appli
12	155	25.1	6464	US-08-400-159-5	Sequence 5, Appli

13	155	25.1	6464	3	US-08-611-729A-5	Sequence 5, Appli
14	154.5	25.0	3582	2	US-08-400-159-9	Sequence 9, Appli
15	154.5	25.0	3582	3	US-08-611-729A-9	Sequence 9, Appli
16	152.5	24.7	7286	4	US-08-793-273C-1	Sequence 1, Appli
17	152.5	24.7	7286	5	PCT-US95-11684-1	Sequence 1, Appli
18	146.5	23.7	1556	1	US-08-597-545-3	Sequence 3, Appli
19	146.5	23.7	1556	1	US-08-457-135-3	Sequence 3, Appli
20	146.5	23.7	2857	3	US-08-981-392-4	Sequence 4, Appli
21	146	23.7	8091	4	US-09-230-652-1	Sequence 1, Appli
22	146	23.7	8257	4	US-09-484-970B-65	Sequence 65, Appl
23	144	23.3	716	4	US-091423-753-11	Sequence 11, Appl
24	144	23.3	2508	3	US-08-981-392-1	Sequence 1, Appli
25	144	23.3	2663	4	US-09-068-740A-8	Sequence 8, Appli
26	144	23.3	2663	4	US-09-423-753-8	Sequence 8, Appli
27	144	23.3	2883	3	US-08-981-392-3	Sequence 3, Appli
28	143.5	23.3	4315	3	US-08-882-046-3	Sequence 3, Appli
29	143.5	23.3	4464	2	US-08-400-159-7	Sequence 7, Appli
30	143.5	23.3	4483	3	US-08-611-729A-7	Sequence 7, Appli
31	142.5	23.1	1981	3	US-08-981-392-26	Sequence 26, Appl
32	142.5	23.1	2005	4	US-09-635-872A-18	Sequence 18, Appl
33	142.5	23.1	2005	4	US-09-636-077A-18	Sequence 18, Appl
34	142.5	23.1	2005	4	US-09-636-060C-18	Sequence 18, Appl
35	142.5	23.1	2005	4	US-09-986-552-18	Sequence 18, Appl
36	142.5	23.1	4080	4	US-09-016-434-1326	Sequence 1326, Ap
37	142.5	23.1	5645	4	US-09-023-655-1319	Sequence 1319, Ap
38	142.5	23.1	5681	4	US-09-919-172-58	Sequence 58, Appl
39	142.5	23.1	5933	4	US-09-919-172-23	Sequence 23, Appl
40	141.5	22.9	2899	3	US-08-981-392-24	Sequence 24, Appl
41	140.5	22.8	3955	3	US-09-214-278-4	Sequence 4, Appli
42	140.5	22.8	3955	4	US-09-855-722-4	Sequence 4, Appli
43	140	22.7	2206	4	US-09-907-794A-3	Sequence 3, Appli
44	140	22.7	2206	4	US-09-905-125A-3	Sequence 3, Appli
45	140	22.7	2206	4	US-09-902-775A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-833-381-1072
; Sequence 1072, Application US/098333381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1072
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(447)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1072

Alignment Scores:
Pred. No.: 1.36e-36 Length: 447
Score: 430.50 Matches: 72
Percent Similarity: 88.89% Conservative: 0
Best Local Similarity: 88.89% Mismatches: 0
Query Match: 69.77% Indels: 9
DB: 4 Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x US-09-833-381-1072 (1-447)

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||||

23	TGCCGGCAGCGCAAGTCGAGTGCAGCCCTGGCTGGATGGCGAACACTGCACCATC---	79
41	HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg	60
80	-----GAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA	115
61	CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly	80
116	TGTACCTTAGACCTGAATGGTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC	175
81	CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu	100
176	TGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAAAGACAATGATGGAGATGGCCTG	235
101	Val 101	
236	GTG 238	
SULT 2		
-08-400-159-1		
Sequence 1, Application US/08400159		
Patent No. 5869282		
GENERAL INFORMATION:		
APPLICANT: Ish-Horowicz, David		
APPLICANT: Henrique, Domingos M.P.		
APPLICANT: Lewis, Julian H.		
APPLICANT: Myat, Anna M.		
APPLICANT: Fleming, Robert J.		
APPLICANT: Artavanis-Tsakonas, Spyridon		
APPLICANT: Mann, Robert S.		
APPLICANT: Gray, Grace E.		
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE		
TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON		
NUMBER OF SEQUENCES: 20		
CORRESPONDENCE ADDRESS:		
ADDRESSEE: Pennie & Edmonds		
STREET: 1155 Avenue of the Americas		
CITY: New York		
STATE: New York		
COUNTRY: USA		
ZIP: 10036-2711		
COMPUTER READABLE FORM:		
MEDIUM TYPE: Floppy disk		
COMPUTER: IBM PC compatible		
OPERATING SYSTEM: PC-DOS/MS-DOS		
SOFTWARE: PatentIn Release #1.0, Version #1.30		
CURRENT APPLICATION DATA:		
APPLICATION NUMBER: US/08/400,159		
FILING DATE: 07-MAR-1995		
CLASSIFICATION: 435		
ATTORNEY/AGENT INFORMATION:		
NAME: Misrock, S. Leslie		
REGISTRATION NUMBER: 18,872		
REFERENCE/DOCKET NUMBER: 7326-029		
TELECOMMUNICATION INFORMATION:		
TELEPHONE: (212) 790-9090		
TELEFAX: (212) 869-9741/8864		
TELEX: 66141 PENNIE		
INFORMATION FOR SEQ ID NO: 1:		
SEQUENCE CHARACTERISTICS:		
LENGTH: 5561 base pairs		
TYPE: nucleic acid		
STRANDEDNESS: double		
TOPOLOGY: unknown		
MOLECULE TYPE: cDNA		
FEATURE:		
NAME/KEY: CDS		
LOCATION: 442..4656		
US-08-400-159-1		
Alignment Scores:		
red. No.:		
core:		
6.48e-08		
169.50		
Length:		
Matches:		
5561		
35		

Percent Similarity:	45.16%	Conservative:	7
Best Local Similarity:	37.63%	Mismatches:	34
Query Match:	27.47%	Indels:	17
DB:	2	Gaps:	6
US-10-029-020-14_COPY_750_850 (1-101) x US-08-400-159-1 (1-5561)			
QY	1	TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCys---AlaGluHisGly	19
Db	1261	TGGCAGGGCGTCAACTGCGAGGAGGCCATATGCAAGCGGGCTGCGACCCCGTCCACGGC	1320
QY	20	ThrCys---ArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr	38
Db	1321	AAGTCCGATCGTCCGGGGGAATGCAATGCAGATGCAGACCCGGTGGTCCATTGTGCAAC	1380
QY	39	IleAlaHisTyrLeuAspArgValValLysGluGlyCysPro---GlyLeuCysAsnGly	57
Db	1381	-----GAGTGCATGGTCTATCCCGCTGCAAGCATGGTTCTCTGCAACGGC	1425
QY	58	AsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArg	77
Db	1426	-----AGCGCTGGAATCGTGTGCGACACCAACTGGGT	1461
QY	78	GlyAlaGlyCysAspThrSerMetGluThrAlaCysGly	90
Db	1462	GGCATATTGTGCGATCAAGATTTA---AATTCTGCGGC	1497
RESULT 3			
US-08-611-729A-1			
; Sequence 1, Application US/08611729A			
; Patent No. 6004924			
; GENERAL INFORMATION:			
; APPLICANT: Ish-Horowicz, David			
; APPLICANT: Henrique, Domingos M.P.			
; APPLICANT: Lewis, Julian H.			
; APPLICANT: Myat, Anna M.			
; APPLICANT: Fleming, Robert J.			
; APPLICANT: Artavanis-Tsakonas, Spyridon			
; APPLICANT: Mann, Robert S.			
; APPLICANT: Gray, Grace E.			
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE			
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON			
; NUMBER OF SEQUENCES: 20			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Pennie & Edmonds			
; STREET: 1155 Avenue of the Americas			
; CITY: New York			
; STATE: New York			
; COUNTRY: U.S.A.			
; ZIP: 10036-2711			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/611,729A			
; FILING DATE: 06-MAR-1996			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Misrock, S. Leslie			
; REGISTRATION NUMBER: 18,872			
; REFERENCE/DOCKET NUMBER: 7326-037			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (212) 790-9090			
; TELEFAX: (212) 869-9741/8864			
; TELEX: 66141 PENNIE			
; INFORMATION FOR SEQ ID NO: 1:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 5561 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: unknown			
; MOLECULE TYPE: cDNA			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: 442..4656			
; US-08-400-159-1			
Alignment Scores:			
red. No.:			
core:			
6.48e-08			
169.50			
Length:			
Matches:			
5561			
35			

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
:::|||||
Db 1875 TTCATGGGGAGGACTGCGGGAGCTGTCTCTGCCAATGACTGCCACCAGCAGCGGCGC 1934

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
:::|||||
Db 1935 TGCCTCATGGCGCTGCTGTGCCACGAGGGCTTCACTGGGAAGACTGCCGG----- 1988

QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
:::|||||
Db 1989 -----GAACGGTCTCTGCCCAATGACTGCAACAACGTTGGGCGC 2027

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
:::|||||
Db 2028 TGT-----GTCGAGGGA---CGGTGTGTCTGTGAGGAAGTTACATGGGGATCGAC 2075

QY 81 Cys---AspThrSerMetGluThr 87
:::|||||
Db 2076 TGTCTGATGTCTCTCTCCCAACG 2099

RESULT 6

PCT-US95-11684-3
; Sequence 3, Application PC/TUS9511684
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING SAME
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11684
; FILING DATE: 14-SEP-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,359
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Logan, April C.
; REGISTRATION NUMBER: 33,950
; REFERENCE/DOCKET NUMBER: BEC0019P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 309..5740
PCT-US95-11684-3

Alignment Scores:
Pred. No.: 7.09e-07 Length: 6049
Score: 160.00 Matches: 33
Percent Similarity: 48.86% Conservative: 10
Best Local Similarity: 37.50% Mismatches: 31

Query Match: 25.93% Indels: 14
DB: 5 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x PCT-US95-11684-3 (1-6049)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
:::|||||
Db 1875 TTCATGGGGAGGACTGCGGGAGCTGTCTCTGCCAATGACTGCCACCAGCAGCGGCGC 1934

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
:::|||||
Db 1935 TGCCTCATGGCGCTGCTGTGCCACGAGGGCTTCACTGGGAAGACTGCCGG----- 1988

QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
:::|||||
Db 1989 -----GAACGGTCTCTGCCCAATGACTGCAACAACGTTGGGCGC 2027

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
:::|||||
Db 2028 TGT-----GTCGAGGGA---CGGTGTGTCTGTGAGGAAGTTACATGGGGATCGAC 2075

QY 81 Cys---AspThrSerMetGluThr 87
:::|||||
Db 2076 TGTCTGATGTGTCTCTCCCAACG 2099

RESULT 7

US-09-214-278-6
; Sequence 6, Application US/09214278
; Patent No. 6291210
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/214,278
; CURRENT FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)..(4062)
; NAME/KEY: sig_peptide
; LOCATION: (409)..(501)
; NAME/KEY: mat_peptide
; LOCATION: (502)..(4062)
US-09-214-278-6

Alignment Scores:
Pred. No.: 1.48e-06 Length: 4208
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 3 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-09-214-278-6 (1-4208)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
:::|||||
Db 1078 TGGATGGGCCCCGAATGTAACAGAGCTATTGCCGACAAGGCTGCAGTCCCTAAGCATGGG 1137

QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
:::|||||
Db 1138 TCTTGCAAACTCCAGGTGACTGCAGGTGCCAGTACGGCTGGCAAGGCTGTACTGT--- 1194

QY 39 IleAlaHisTyrLeuAspArgValVallysGlu---GlyCys---ProGlyLeuCysAsn 56
:::|||||
Db 1195 -----GATAAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAT 1239

QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1240 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1272
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1273 GCGGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1311

RESULT 8
US-09-068-740A-10
; Sequence 10, Application US/09068740A
; Patent No. 6337387
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; APPLICANT: ITOH, AKIRA
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8447
; CURRENT APPLICATION NUMBER: US/09/068,740A
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: JP 7-299611
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: JP 7-311811
; PRIOR FILING DATE: 1995-11-30
; PRIOR APPLICATION NUMBER: PCT/JP96/03356
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)..(4062)
US-09-068-740A-10

Alignment Scores:
Pred. No.: 1.48e-06 Length: 4208
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 4 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-09-068-740A-10 (1-4208)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1078 TGGATGGGCCCCGGAATGTAACAGAGCTATTTGCCGACAAGGCTGCAGTCCTAAGCATGGG 1137
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1138 TCTTGCAAACCTCCAGGTGACTGCAGGTGCCAGTACGGCTGGCAAGGCTGTACTGT--- 1194
QY 39 IleAlaHisTyrLeuAspArgValValLysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1195 -----GATAAGTGCATCCACACCCCGGATCGTCCACGGCATCTGTAAT 1239
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1240 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1272
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1273 GCGGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1311

RESULT 9
US-09-855-722-6
; Sequence 6, Application US/09855722
; Patent No. 6638741
; GENERAL INFORMATION:
; APPLICANT: Sakano, Seiji
; APPLICANT: Itoh, Akira

; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8576
; CURRENT APPLICATION NUMBER: US/09/855,722
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/214,278
; PRIOR FILING DATE: 1999-01-26
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 4208
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (409)..(4062)
; NAME/KEY: sig_peptide
; LOCATION: (409)..(501)
; NAME/KEY: mat_peptide
; LOCATION: (502)..(4062)
US-09-855-722-6
Alignment Scores:
Pred. No.: 1.48e-06 Length: 4208
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 4 Gaps: 7
US-10-029-020-14_COPY_750_850 (1-101) x US-09-855-722-6 (1-4208)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1078 TGGATGGGCCCCGGAATGTAACAGAGCTATTTGCCGACAAGGCTGCAGTCCTAAGCATGGG 1137
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1138 TCTTGCAAACCTCCAGGTGACTGCAGGTGCCAGTACGGCTGGCAAGGCTGTACTGT--- 1194
QY 39 IleAlaHisTyrLeuAspArgValValLysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1195 -----GATAAGTGCATCCACACCCCGGATCGTCCACGGCATCTGTAAT 1239
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1240 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1272
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1273 GCGGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1311
RESULT 10
US-09-199-865-2
; Sequence 2, Application US/09199865
; Patent No. 6433138
; GENERAL INFORMATION:
; APPLICANT: Zimrin, Ann B.
; APPLICANT: Maciag, Thomas
; APPLICANT: Wong, Michael K.K.
; APPLICANT: Pepper, Michael S.
; APPLICANT: Montesano, Roberto
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC METHODS AND COMPOSITIONS
; FILE REFERENCE: 0036-1U1
; CURRENT APPLICATION NUMBER: US/09/199,865
; CURRENT FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: 60/018,841
; EARLIER FILING DATE: 1996-05-31
; EARLIER APPLICATION NUMBER: PCT/US97/09407
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

```

; LENGTH: 5458
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-199-865-2

Alignment Scores:
Pred. No.: 2.07e-06 Length: 5458
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservatives: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 4 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-09-199-865-2 (1-5458)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 640 TGGATGGGCCCCGAATGTAACAGAGCTATTGCGCACAAGGCTGCAGTCCTTAAGCATGGG 699

QY 20 ThrCysArg---AspGlylysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 700 TCTTGCAAACTCCAGGTGACTGCAGGTGCCAGTACGGCTGGCAAGGCCTGTACTGT--- 756

QY 39 IleAlaHisTyrLeuAspArgValVallysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 757 -----GATAAGTGCATCCACACCCCGGATGCGTCCACGGCATCTGTAAT 801

QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 802 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 834

QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 835 GCGGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 873

RESULT 11
US-08-882-046-1
; Sequence 1, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5590 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 414..4068
US-08-882-046-1

Alignment Scores:
Pred. No.: 2.13e-06 Length: 5590
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservatives: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 3 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-08-882-046-1 (1-5590)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1083 TGGATGGGCCCCGAATGTAACAGAGCTATTGCGCACAAGGCTGCAGTCCTTAAGCATGGG 1142

QY 20 ThrCysArg---AspGlylysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1143 TCTTGCAAACTCCAGGTGACTGCAGGTGCCAGTATGGCTGGCAAGGCCTGTACTGT--- 1199

QY 39 IleAlaHisTyrLeuAspArgValVallysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1200 -----GATAAGTGCATCCACACCCCGGATGCGTCCACGGCATCTGTAAT 1244

QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1245 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1277

RESULT 12
US-08-400-159-5
; Sequence 5, Application US/08400159
; Patent No. 5869282
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie

```

REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6464 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 371..4027
US-08-400-159-5

Alignment Scores:
Pred. No.: 2.57e-06 Length: 6464
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 2 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-08-400-159-5 (1-6464)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1040 TGGATGGGCCCCGAATGTAACAGAGCTATTGGCCGACAAAGCTGCAGTCCTAAGCATGGG 1099
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1100 TCTTGCAAACTCCCAGGTGACTGCAGGTGCCAGTACGGCTGGCAAGGCCTGTACTGT--- 1156
QY 39 IleAlaHisTyrLeuAspArgValVallysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1157 -----GATAAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAT 1201
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1202 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1234
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1235 GCGGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1273

RESULT 13
US-08-611-729A-5
Sequence 5, Application US/08611729A
Patent No. 6004924
GENERAL INFORMATION:
APPLICANT: Ish-Horowicz, David
APPLICANT: Henrique, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,729A
FILING DATE: 06-MAR-1996
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-037
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 6464 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 371..4024
US-08-611-729A-5

Alignment Scores:
Pred. No.: 2.57e-06 Length: 6464
Score: 155.00 Matches: 31
Percent Similarity: 46.81% Conservative: 13
Best Local Similarity: 32.98% Mismatches: 30
Query Match: 25.12% Indels: 20
DB: 3 Gaps: 7

US-10-029-020-14_COPY_750_850 (1-101) x US-08-611-729A-5 (1-6464)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAla---GluHisGly 19
Db 1040 TGGATGGGCCCCGAATGTAACAGAGCTATTGGCCGACAAAGCTGCAGTCCTAAGCATGGG 1099
QY 20 ThrCysArg---AspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThr 38
Db 1100 TCTTGCAAACTCCCAGGTGACTGCAGGTGCCAGTACGGCTGGCAAGGCCTGTACTGT--- 1156
QY 39 IleAlaHisTyrLeuAspArgValVallysGlu---GlyCys---ProGlyLeuCysAsn 56
Db 1157 -----GATAAGTGCATCCACACCCGGGATGCGTCCACGGCATCTGTAAT 1201
QY 57 GlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrp 76
Db 1202 -----GAGCCCTGGCAGTGCCTCTGTGAGACCAACTGG 1234
QY 77 ArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGly 90
Db 1235 GCGGCCAGCTCTGTGACAAAGATCTC---AATTACTGTGGG 1273

RESULT 14
US-08-400-159-9
Sequence 9, Application US/08400159
Patent No. 5869282
GENERAL INFORMATION:
APPLICANT: Ish-Horowicz, David
APPLICANT: Henrique, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,159
; FILING DATE: 07-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3582
; US-08-400-159-9

Alignment Scores:
Pred. No.: 1.35e-06 Length: 3582
Score: 154.50 Matches: 33
Percent Similarity: 41.80% Conservative: 18
Best Local Similarity: 27.05% Mismatches: 40
Query Match: 25.04% Indels: 31
DB: 2 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x US-08-400-159-9 (1-3582)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAla-----CysHisProArgCysAlaGluHis 18
Db 1246 TGGTCTGGCCACAACCTGTGATATAATAATTAATGATTGTCGTGGACAATGTCAGAATGGA 1305
QY 19 GlyThrCysArgAsp-----GlyLysCysGluCysSerProGlyTrpAsnGly 34
Db 1306 GGATCCTCTCGGGACTTGGTTAATGGTTATCGGTGCATCTGTTCACTGGCTATGCAGGA 1365
QY 35 GluHisCysThrIleAlaHisTyrLeuAspArgValVallLysGluGlyCysProGlyLeu 54
Db 1366 GATCACTGT-----GAGAAAGACATCAATGATGTCGAAGTAACCCCT 1407
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 1408 TGCATGATGGGGGTCACTGCCAGGATGAATCAATGGATTCCCAATGCTGTGTCTCTGCT 1467
QY 75 GlyTrpArgGlyAlaGlyCys-----81
Db 1468 GGTTTCTCAGGAACCTCTGTTCAGCTGGATATAGACTACTGTGAGCCCAACCCCTTGCCAG 1527
QY 82 -----AspThrSerMetGluThrAlaCysGlyAspSerLysAspAsn 95
Db 1528 AACGGTGCCAGTGCTTCAATCTTGCTATGGACTATTCTGTAACTGCCCTGAAGATTAC 1587
QY 96 AspGly 97
Db 1588 GAAGGC 1593

RESULT 15
US-08-611-729A-9
; Sequence 9, Application US/08611729A
; Patent No. 6004924
; GENERAL INFORMATION:
; APPLICANT: Ish-Horowicz, David
; APPLICANT: Henrique, Domingos M.P.
; APPLICANT: Lewis, Julian H.
; APPLICANT: Myat, Anna M.
; APPLICANT: Fleming, Robert J.
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Mann, Robert S.
; APPLICANT: Gray, Grace E.
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF THE
; TITLE OF INVENTION: SERRATE GENE AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/611,729A
; FILING DATE: 06-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..3579
; US-08-611-729A-9
Alignment Scores:
Pred. No.: 1.35e-06 Length: 3582
Score: 154.50 Matches: 33
Percent Similarity: 41.80% Conservative: 18
Best Local Similarity: 27.05% Mismatches: 40
Query Match: 25.04% Indels: 31
DB: 3 Gaps: 4
US-10-029-020-14_COPY_750_850 (1-101) x US-08-611-729A-9 (1-3582)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAla-----CysHisProArgCysAlaGluHis 18
Db 1246 TGGTCTGGCCACAACCTGTGATATAATAATTAATGATTGTCGTGGACAATGTCAGAATGGA 1305
QY 19 GlyThrCysArgAsp-----GlyLysCysGluCysSerProGlyTrpAsnGly 34
Db 1306 GGATCCTCTCGGGACTTGGTTAATGGTTATCGGTGCATCTGTTCACTGGCTATGCAGGA 1365
QY 35 GluHisCysThrIleAlaHisTyrLeuAspArgValVallLysGluGlyCysProGlyLeu 54

```
Db 1366 GATCACTGT-----GAGAAAGACATCAATGAATGTGCAAGTAACCCCT 1407
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTyrHisCysValCysGlnLeu 74
Db 1408 TGCATGAATGGGGTCACTGCTGAGGATGAAATCAATGGATTCCAATGTCTGTCTGCTGCT 1467
QY 75 GlyTyrArgGlyAlaGlyCys----- 81
Db 1468 GGTTTCTCAGGAAACCTCTGTGCTGAGCTGGATATAGACTACTGTGAGCCAAACCCCTTGCCAG 1527
QY 82 -----AspThrSerMetGluThrAlaCysGlyAspSerLysAspAsn 95
Db 1528 AACGGTGCCAGTGCTTCAATCTTGCTATGGACTATTTCTGTAACTGCCCTGAAGATTAC 1587
QY 96 AspGly 97
Db 1588 GAAGGC 1593
```

Search completed: August 14, 2004, 19:40:00
Job time : 44.8807 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 1184.51 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WMGAAACDQACHPRCAEHGT.....DTSMETACGDSKNDGDGLV 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@CGN_1_1_13135@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	543.5	88.1	610	10	BB633431	BB633431 BB633431
2	543.5	88.1	1284	11	AK039472	AK039472 Mus muscu
3	537.5	87.1	464	10	BE671538	BE671538 7e53f01.x
4	470.5	76.3	435	14	CB793108	CB793108 AMGNNUC:N
5	454	73.6	462	10	BE646762	BE646762 UI-M-BH1-
6	439	71.2	718	13	BU341069	BU341069 603406859
7	410.5	66.5	847	14	CK017330	CK017330 AGENCOURT
8	323	52.4	314	9	AL710634	AL710634 DKFZp886G
9	261.5	42.4	492	14	CD163026	CD163026 ML1-0081T
10	254.5	41.2	501	14	CD202234	CD202234 MS1-0131U
11	252.5	40.9	633	10	BB665486	BB665486 BB665486
12	252.5	40.9	683	10	BB665259	BB665259 BB665259
13	250	40.5	1376	11	AK053790	AK053790 Mus muscu
14	250	40.5	3190	29	AY413477	AY413477 Mus muscu
15	250	40.5	6246	29	AY413475	AY413475 Homo sapi
16	247	40.0	2729	11	AK031173	AK031173 Mus muscu
17	224	36.3	550	28	AZ393761	AZ393761 IM0156P20
18	223.5	36.2	511	10	BE650761	BE650761 UI-M-BH3-
19	223.5	36.2	2627	11	AK050784	AK050784 Mus muscu
20	216	35.0	1040	29	CNS006TS	AL065915 Drosophil
21	214.5	34.8	1101	29	CNS00056	AL050564 Drosophil
22	210	34.0	252	10	BE088902	BE088902 CM2-BT069
23	210	34.0	252	10	BE088973	BE088973 CM2-BT069
24	206	33.4	1101	29	CNS000GVU	AL072610 Drosophil
25	202.5	32.8	376	14	CD118097	CD118097 ME1-0044T
26	200	32.4	659	13	BQ284539	BQ284539 fz88005.x
27	198	32.1	1023	29	CNS006NW	AL065602 Drosophil
28	171	27.7	724	13	BU387957	BU387957 603583120
29	171	27.7	870	29	CNS01UF5	AL167738 Tetraodon
30	171	27.7	2716	11	AK031198	AK031198 Mus muscu
31	170.5	27.6	395	13	BW188423	BW188423 BW188423
32	170.5	27.6	572	9	AL588592	AL588592 AL588592
33	169.5	27.5	589	12	BI358014	BI358014 RE44666.5
34	169.5	27.5	609	12	BI356067	BI356067 RE42104.5
35	169.5	27.5	733	13	BU110388	BU110388 603002245
36	169.5	27.5	737	13	BU325086	BU325086 603492835
37	169.5	27.5	751	13	BU274274	BU274274 603530626
38	169.5	27.5	790	13	BU273516	BU273516 603533089
39	169.5	27.5	805	13	BU284708	BU284708 603601739
40	168.5	27.3	626	12	BI393390	BI393390 pgpln.pk0
41	167.5	27.1	659	13	BQ284142	BQ284142 fz88d05.y
42	167.5	27.1	805	13	BU340855	BU340855 603518278
43	166.5	27.0	611	13	BU462653	BU462653 603366851
44	165.5	26.8	675	13	BU403903	BU403903 604138914
45	163	26.4	755	13	BU408094	BU408094 603481266

ALIGNMENTS

RESULT 1 BB633431
LOCUS BB633431
DEFINITION BB633431 RIKEN full-length enriched, adult male spinal cord Mus
musculus cDNA clone A330048C04 5', mRNA sequence.
ACCESSION BB633431
VERSION BB633431.1 GI:16469876
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 610)

AUTHORS

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

TITLE
JOURNAL
COMMENT

Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

e mouse tissues.

FEATURES

source

Location/Qualifiers

1. .610
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A330048C04"
/sex="male"
/tissue_type="spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male spinal
cord"

/note="Site 1: Sall; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAAGGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGATTCTCGATTATTAATTAATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

ORIGIN

Alignment Scores:

Pred. No.: 2.5e-37 Length: 610
Score: 543.50 Matches: 89
Percent Similarity: 88.12% Conservative: 0
Best Local Similarity: 88.12% Mismatches: 3
Query Match: 88.09% Indels: 9
DB: 10 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x BB633431 (1-610)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 167 TGGATGGGGCCGCATGCGACCAACGGGCGCTGCCACCCACCGCTGTGCAGAACACGGGACC 226
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 227 TGCCGGGACGGCAAGTGCGAATGCAGCCCCCGCTGGATGGAGAGCACTGCACCATC--- 283
QY 41 HisTyrLeuAspArgValLlysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 284 -----GAGGGCTGTCTCTGCTGCAATGGAAATGGCAGA 319
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 320 TGTACCCCTGGACCTGAATGGGTGGCACTGTGTCTGCCAGCTGGGCTGGCGAGGACTGGC 379
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnGlyAspGlyLeu 100
Db 380 TGGACACATCCATCGAAACGGGCTGTGGAGACGGCAAGGACGACGAGATGGCTTG 439
QY 101 Val 101
Db 440 GTG 442

RESULT 2

AK039472

LOCUS

DEFINITION

AK039472 1284 bp mRNA linear HTC 19-SEP-2003
Mus musculus adult male spinal cord cDNA, RIKEN full-length
enriched library, clone:A330048C04 product:neuregulin 1, full
insert sequence.

ACCESSION

AK039472.1 GI:26087188

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE

1 Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

99279253

MEDLINE

10349636

PUBMED

11042159

REFERENCE

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

20499374

MEDLINE

11042159

PUBMED

11042159

REFERENCE

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,

Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,

Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,

Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,

Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

11042159

MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 1284)
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source Location/Qualifiers
1. .1284
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A330048C04"
/db_xref="MGI:2403942"
/db_xref="taxon:10090"
/clone="A330048C04"
/sex="male"
/tissue type="spinal cord"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature 1. .1284
/note="neuregulin 1 (MGD|MGI:96083, MGD|MGI:1345182, GB|NM_011858, evidence: BLASTN, 99%, match=186)"
ORIGIN
Alignment Scores:
Pred. No.: 7.03e-37 Length: 1284
Score: 543.50 Matches: 89
Percent Similarity: 88.12% Conservative: 0
Best Local Similarity: 88.12% Mismatches: 3
Query Match: 88.09% Indels: 9
DB: 11 Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x AK039472 (1-1284)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 167 TGGATGGGGCCGCATGCGACCAACGGGGCCTGCCACCCACGCTGTGCAGAACACGGGACC 226

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 227 TGGCGGACGGCAAGTGCGAATGCAGCCCGGCTGGAATGGAGAGCACTGCACCATC--- 283
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 284 -----GAGGGCTGTCTGGCTGTGCAATGGAAATGGCAGA 319
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 320 TGTACCTTGGACCTGAATGGTGGCACTGTGTCTGCCAGCTGGGCTGGCGAGGACTGGC 379
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 380 TCGACACATCCATCGAAACGGGCTGTGGAGACGGCAAGCAACGACGAGATGGCTTG 439
QY 101 Val 101
Db 440 GTG 442
RESULT 3
BE671538
LOCUS BE671538 464 bp mRNA linear EST 08-SEP-2000
DEFINITION 7e53f01.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3286201 3, similar to TR:O70465 O70465 DOC4. ;, mRNA sequence.
ACCESSION BE671538
VERSION BE671538.1 GI:10032079
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 464)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco.
FEATURES
source Location/Qualifiers
1. .464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3286201"
/tissue type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI-CGAP_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."
ORIGIN
Alignment Scores:
Pred. No.: 5.67e-37 Length: 464
Score: 537.50 Matches: 88
Percent Similarity: 89.80% Conservative: 0

Best Local Similarity: 89.80% Mismatches: 1
Query Match: 87.12% Indels: 9
DB: 10 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x BE671538 (1-464)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 197 TGGATGGGGCAGCCTGCGACGAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACC 256

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 257 TGCCGCGACGGCAAGTCGAGTCAGCCCTGGCTGGAATGGCGAACACTGCACCATC--- 313

QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 314 -----GAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 349

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 350 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGGC 409

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAsp 98
Db 410 TGTGACACTTCCATGGAGACTGCTGCGGTGACCAGAAACACAATGATGGAGAT 463

RESULT 4
CB793108
LOCUS
DEFINITION AMGNNUC:NRHY5-00339-E12-A W Rat hypothalamus (10471) Rattus norvegicus cDNA clone nrhy5-00339-e12 5', mRNA sequence.
ACCESSION CB793108.1 GI:29881501
VERSION
KEYWORDS
SOURCE
ORGANISM Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 435)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00339 row: e column: 12.

FEATURES
source
1. .435
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrhy5-00339-e12"
/clone_lib="W Rat hypothalamus (10471)"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; W Rat hypothalamus adult female Wistar rat avg. insert size 2.3 kb fraction 6 and 7"

ORIGIN
Alignment Scores:
Pred. No.: 3.28e-31 Length: 435
Score: 470.50 Matches: 73
Percent Similarity: 83.00% Conservative: 10
Best Local Similarity: 73.00% Mismatches: 16
Query Match: 76.26% Indels: 1
DB: 14 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x CB793108 (1-435)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 60 TGGGTAGACCAACATGTGAGGACGTTCTCTGCTACTCTCCTGCTGCTGAGCATGGTCAG 119

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 120 TGCAGAGACGGAAATGTAGTGTAGCCCTGGATGGGAGGCGCACCACTGCCACAATTGCT 179

QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 180 CACTACTTAGAT---GCTGTTTCGAGATGGCTGTCCAGGCTCTGCTTTGGAAATGGACGA 236

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 237 TGTACCCTGGATCAAAATGGTTGGCACTGTGTGTGTCAGGTGGTTGGAGTGGAACAGGT 296

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 297 TGCAACATTGTCATGGAATGCTTTGTGGCGACAACCTTGGACAATGATGGAGATGGTTTG 356

RESULT 5
BE646762
LOCUS
DEFINITION UI-M-BH1-akr-d-12-0-UI.r1 NIH BMAP_M_S2 Mus musculus cDNA clone UI-M-BH1-akr-d-12-0-UI 5', mRNA sequence.
ACCESSION BE646762
VERSION BE646762.1 GI:9972517
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 462)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements
Seq primer: M13 Reverse.

FEATURES
source
1. .462
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH1-akr-d-12-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH BMAP_M_S2"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH BMAP_M_S2 library is a subtracted library derived from NIH_BMAP_M_S1, which in turn is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 5,000 clones from the NIH_BMAP_M_S1 library and a pool of 2,000 clones obtained from non-normalized and normalized mouse brain spinal cord libraries."

ORIGIN

Alignment Scores:
Pred. No.: 9.58e-30 Length: 462
Score: 454.00 Matches: 71
Percent Similarity: 77.23% Conservative: 7
Best Local Similarity: 70.30% Mismatches: 23
Query Match: 73.58% Indels: 0
DB: 10 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x BE646762 (1-462)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 21 TGGACAGGCGCAGCTTGTGACCAGCGGTGTGCCACCCCGCTGCATTGAGCACGGGACC 80
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 81 TGTAAAGATGGCAATGTGAATGCCGAGAGGGCTGGAAATGGTGAACACTGCACCATGGT 140
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 141 AGGCAAAACGGCAGGCACCGAAACACAGATGGCTGCCCTGATTGTGCAACGGTAAACGGGAGA 200
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 201 TGCACACTGGGTTCAGAACAGCTGGCAGTGTGTCTGCAGACCGGCTGGAGAGGGCTGGA 260
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 261 TGCAACGTTGCCATGGAAACCTCCTGCGCTGATACACAGGATAATGAGGAGATGGCCTG 320
QY 101 Val 101
Db 321 GTG 323

RESULT 6
BU341069 718 bp mRNA linear EST 28-NOV-2002
LOCUS 603406859F1 CSEQCHN67 Gallus gallus cdna clone ChEST315e16 5', mRNA
DEFINITION sequence.
ACCESSION BU341069
VERSION BU341069.1 GI:25849070
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 718)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .718
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST315e16"
/dev_stage="16 day embryo"
/lab_host="DH10B"
/clone_lib="CSEQCHN67"

FEATURES
source
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST315e16"
/dev_stage="16 day embryo"
/lab_host="DH10B"
/clone_lib="CSEQCHN67"

/note="Organ: brain; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Alignment Scores:
Pred. No.: 3.51e-28 Length: 718
Score: 439.00 Matches: 70
Percent Similarity: 75.49% Conservative: 7
Best Local Similarity: 68.63% Mismatches: 24
Query Match: 71.15% Indels: 1
DB: 13 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x BU341069 (1-718)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 31 TGGACAGGAGTGGCGTGTGACCAGCGTGTGTGTCATCCCGGTGTACAGAGCACGGAAC 90
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 91 TGTAAAGATGGGAAATGTGAATGCAGAGAGGGCTGGAAATGGGAGCACTGCACCATGGT 150
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 151 AGGCAAAACGACAGGCACCGAAACAGATGGCTGCCCTGACTTGTGCAATGGCAACGGGAGG 210
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 211 TGCACGCTGGGGCCAGAACAGCTGGCAGTGTGTCTGCCAGACCGGCTGGAGAGGGCTGGA 270
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp-AsnAspGlyAspGlyLe 100
Db 271 TGCAACGTTGCCATGGAAACCTCCTGTGCCGATAACAAGGATAAACGAGGAGATGGCTT 330
QY 100 uVal 101
Db 331 GGTT 334

RESULT 7
CK017330 847 bp mRNA linear EST 26-NOV-2003
LOCUS AGENCOURT_16543817 NIH_ZGC_10 Danio rerio cdna clone IMAGE:7044835
DEFINITION 5', mRNA sequence.
ACCESSION CK017330
VERSION CK017330.1 GI:38543254
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 847)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Len Zon, Harvard

cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14806 row: 1 column: 17
High quality sequence stop: 585.
Location/Qualifiers
1. .847
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7044835"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/note="Vector: pExpress1; Site_1: NotI; Site_2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC_7). Library was constructed by Open Biosystems
(Huntsville, AL)."

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 1.3e-25 Length: 847
Score: 410.50 Matches: 64
Percent Similarity: 71.29% Conservative: 8
Best Local Similarity: 63.37% Mismatches: 20
Query Match: 66.53% Indels: 9
DB: 14 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x CK017330 (1-847)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 209 TGGGAGGGAACCACTTGTGACAAACAGCCCTGCCATCCCAATCTGTGAAGCATGGAGAA 268
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 269 TGCAGGGACGGGCAGTGTCTGTCTGAGCCTGGTTGGAGGAGAGCACTGCACATT--- 325
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 326 -----GATGCCTGCCAGGGTTGTGTAAACGGTAATGGACGC 361
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 362 TGCACCTGGAGCAGACGGGTGGCATTGTGTGCCAGTCTGGATGGAGCGGCCGGC 421
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 422 TGTAAACGTCGTCATGGAAACCGAGTGCAATGACAGTAAAGATAATGACAGCGATGGTCTC 481

QY 101 Val 101
Db 482 ATG 484

RESULT 8
AL710634
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AL710634 314 bp mRNA linear EST 04-SEP-2003
DKFZp686G1170_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DKFZp686G1170_5', mRNA sequence.
AL710634
AL710634.1 GI:19693989
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 314)
Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp686G1170) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .314
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686G1170"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 1.24e-18 Length: 314
Score: 323.00 Matches: 51
Percent Similarity: 72.60% Conservative: 2
Best Local Similarity: 69.86% Mismatches: 20
Query Match: 52.35% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x AL710634 (1-314)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 90 TGGACAGCGCAGCGTGTGACACGCGGTGTGCCACCCCGCTGCATTGACACGGGACC 149
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 150 TGTAAAGATGGCAAATGTGAATCCGAGAGGGCTGGATGGTGAACACTGCACCATTTGT 209
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 210 AGGCAACCGCAGGCACCGAACACAGATGGCTGCCCTGACTGTGTGCAACCGTAAACGGGAGA 269
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGln 73
Db 270 TGCACACTGGGTTCAGAACACAGCTGGCAGTGTGTCTGGCAA 308

RESULT 9
CD163026

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
CD163026 492 bp mRNA linear EST 14-SEP-2003
ML1-0081T-M291-B11-U.G ML1-0081 Schistosoma mansoni cDNA clone
ML1-0081T-M291-B11.G, mRNA sequence.
CD163026
CD163026.1 GI:34699724
EST.
Schistosoma mansoni
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatoida; Schistosomatidae; Schistosoma.

REFERENCE
AUTHORS

1 (bases 1 to 492)
Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,

Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L., Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A., Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A., Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T., Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M., Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
Transcriptome analysis of the acoelomate human parasite Schistosoma mansoni
Nat. Genet. 35 (2), 148-157 (2003)
22879926
Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquimica
Instituto de Quimica - Universidade de Sao Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP, Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjo@iq.usp.br
This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL http://bioinfo.iq.usp.br/schisto/
Plate: ML1-0081T-M291 row: 11 column: B.
Location/Qualifiers
1. 492
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="ML1-0081T-M291-B11.G"
/sex="mixed pool"
/dev_stage="miracidium"
/clone_lib="ML1-0081"
/note="Vector: pGEM T-easy"

ORIGIN
Alignment Scores:
Pred. No.: 4.87e-13 Length: 492
Score: 261.50 Matches: 43
Percent Similarity: 58.25% Conservative: 17
Best Local Similarity: 41.75% Mismatches: 32
Query Match: 42.38% Indels: 11
DB: 14 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x CD163026 (1-492)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 144 TGGACAGGTTCCGATTGTAATACAAAACGATCGGATCTGTTGTTTATAAATGGATAT 203

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 204 TGTACTAATGGTACATGTATTGCAAGTCGGTTGGAATGGAAGTACTGTACATTA--- 260

QY 41 HisTyrLeuAspArgValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 261 -----GATGGTTGCCCAGATCATTTGTAATGGACACCGTCAA 296

QY 61 CysThrLeuAsp-----LeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGly 78
Db 297 TGCACCATGTATCCATTGTCTAGTTCTTATTATTGTCGAATGTTCTCCTGGTTGGTCAGGT 356

QY 79 AlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAsp 98
Db 357 AGTGC GTGTCAACGCAAAATCGAAACAATATGTGATGATGGAATAGATAATGATCATGAT 416

QY 99 GlyLeuVal 101
Db 417 GGTCTTGTT 425

RESULT 10
CD202234
LOCUS
DEFINITION MS1-0131U-A342-A04-U.B MS1-0131 Schistosoma mansoni cDNA clone

MS1-0131U-A342-A04.B, mRNA sequence.
CD202234
CD202234.1 GI:34731943
EST.
Schistosoma mansoni
Schistosoma mansoni
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
1 (bases 1 to 501)
Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M., Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr., Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F., Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L., Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A., Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A., Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T., Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M., Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
Transcriptome analysis of the acoelomate human parasite Schistosoma mansoni
Nat. Genet. 35 (2), 148-157 (2003)
22879926
Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquimica
Instituto de Quimica - Universidade de Sao Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP, Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjo@iq.usp.br
This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL http://bioinfo.iq.usp.br/schisto/
Plate: MS1-0131U-A342 row: 4 column: A.
Location/Qualifiers
1. 501
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="MS1-0131U-A342-A04.B"
/sex="mixed pool"
/dev_stage="schistosomulum"
/lab_host="in vitro culture"
/clone_lib="MS1-0131"
/note="Vector: pGEM T-easy"

ORIGIN
Alignment Scores:
Pred. No.: 2.02e-12 Length: 501
Score: 254.50 Matches: 42
Percent Similarity: 57.28% Conservative: 17
Best Local Similarity: 40.78% Mismatches: 33
Query Match: 41.25% Indels: 11
DB: 14 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x CD202234 (1-501)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 114 TGGACAGGTTCCGATTGTAATACAAAACGATCGGATCTGTTGTTTATAAATGGATAT 173

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 174 TGTACTAATGGTACATGATTTTGCAAAGTCGGTTGGAATGGAAGTACTGTACATTA--- 230

QY 41 HisTyrLeuAspArgValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 231 -----GATGGTTGCCCAGATCATTTGTAATGGACACCGTCAA 266

QY 61 CysThrLeuAsp-----LeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGly 78
Db 267 TGCACCATGTATCCATTGTCTAGTTCTTATTATTGTCGAATGTTCTCCTGGTTGGTCAGGT 326

Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,M. and Hayashizaki,Y.
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K.,
Arakawa,T., Ishii,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Func. Genomics 2 pre, L72-l86 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. .683
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E130115C05"
/tissue_type="eyeball"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate
eyeball"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGCGCAACTCGAGTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATTCGAGTTAATTAATTAATCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Alignment Scores:

Pred. No.:	4.62e-12	Length:	683
Score:	252.50	Matches:	48
Percent Similarity:	69.23%	Conservative:	6
Best Local Similarity:	61.54%	Mismatches:	15
Query Match:	40.92%	Indels:	9
DB:	10	Gaps:	3

US-10-029-020-14_COPY_750_850 (1-101) x BB665259 (1-683)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||||
Db 176 TGGATGGGGCCGATCGACCAACGGGCCTGCCACCACCGCTGTGCAGAACACGGGACC 235
|||||
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||||

Db 236 TGCCGGGACGGCAAGTGCGAATGCAGCCCGGCTGGAATGGAGCACTGCACCATCGGT 295

QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
::: |||||
Db 296 ATG-----GCGGGCGCACGTCTCTGGGGCATTTGGGTGCTCTCAAGACATTTCCAG 346
::: |||||
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGly 78
::: |||||
Db 347 GGCAGCCTGGTTCTC-----CTCCATTGCATG-----TGGATTGGC 382
::: |||||

RESULT 13
AK053790
LOCUS
DEFINITION
AK053790
1376 bp mRNA linear HTC 20-SEP-2003
Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched
library, clone:E130308J18 product:neuregulin 1, full insert
sequence.
AK053790
VERSION
AK053790.1 GI:26095991
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
PUBMED
10349636
2
REFERENCE
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
11042159
3
REFERENCE
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
4
REFERENCE
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
REFERENCE
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1376)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,

gene <1. .>6246
/locus_tag="HCM4903"

ORIGIN

Alignment Scores:

Pred. No.:	1.64e-10	Length:	6246
Score:	250.00	Matches:	39
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	40.52%	Indels:	0
DB:	29	Gaps:	0

US-10-029-020-14_COPY_750_850 (1-101) x AY413475 (1-6246)

QY	1	TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr	20
Db	226	TGGATGGGGGCAGCCTGCGACCCAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACC	285
QY	21	CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle	39
Db	286	TGCCGCGACGGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATC	342

Search completed: August 14, 2004, 18:02:56
Job time : 1193.51 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 13, 2004, 17:03:51 ; Search time 171.323 Seconds
(without alignments)
2504.436 Million cell updates/sec

Title: US-10-029-020-14_COPY_750_850
Perfect score: 617
Sequence: 1 WMGAAACDQACRACHPRCAEHGT.....DTSMETACGDSKNDGDGLV 101

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112215_29265/app_query.fasta_1.3519
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_1_1868@runat_06082004_112215_29265 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	617	100.0	8354	6 ABS52100	Abs52100 Human TEN
2	617	100.0	8645	6 ABS78652	Abs78652 Human cDN
3	579	93.8	8438	6 ABN85378	Abn85378 Human NOV
4	515	83.5	8473	6 ABQ82345	Abq82345 Human NOV
5	515	83.5	8645	6 ABQ82344	Abq82344 Human NOV
6	515	83.5	8675	6 ABQ82343	Abq82343 Human NOV
7	478.5	77.6	8487	6 ABQ82346	Abq82346 Human NOV
8	464.5	75.3	12879	6 ABK92230	Abk92230 Prostate

9	464.5	75.3	13202	4	AAK51828	Human pol
10	454	73.6	9826	5	AAS14085	Human FCT
11	454	73.6	9826	9	ADB32023	Human FCT
12	448.5	72.7	9729	5	AAS14089	Human FCT
13	448.5	72.7	9729	9	ADB32028	Human FCT
14	330.5	53.6	4245	7	AAL60066	Human PC0
15	330.5	53.6	9058	7	ACC72051	BCU0205A
16	330.5	53.6	9695	7	ACC72052	BCU0205B
17	296	48.0	10242	4	ABL29075	Drosophil
18	278	45.1	17131	4	ABL29074	Drosophil
19	277.5	45.0	2569	4	ABL17563	Drosophil
20	169.5	27.5	5552	4	ABL10363	Drosophil
21	169.5	27.5	5561	3	AAZ49099	Drosophil
22	169	27.4	13268	6	ABV94056	Breast ca
23	169	27.4	13857	4	AAI58199	Human pol
24	169	27.4	13857	8	ADB48165	Novel hum
25	169	27.4	100267	6	ABT11032	Human bre
26	160	25.9	6049	2	AAT14548	Cytotacti
27	159.5	25.9	5561	2	AAQ43910	Sequence
28	156	25.3	5353	8	AAL62040	Human cel
29	156	25.3	6696	8	AAL62049	Human cel
30	156	25.3	7560	4	AAC83437	Human ten
31	156	25.3	7560	6	ABL62483	Colon ade
32	156	25.3	7560	7	ABZ77285	Nucleotid
33	156	25.3	7560	9	ADB37427	Human can
34	156	25.3	7560	9	ACF79935	Breast ca
35	156	25.3	7746	6	ABS51834	Novel hum
36	156	25.3	7823	7	ACA04023	CDNA down
37	156	25.3	7823	7	ABX63790	Human cDN
38	155	25.1	4208	2	AAT70175	Prolifera
39	155	25.1	4208	2	AAV15201	Human ser
40	155	25.1	4333	8	ACD06197	Human cDN
41	155	25.1	4855	6	ABK35564	Gene JAG1
42	155	25.1	5458	2	AAV03674	Human Jag
43	155	25.1	5590	2	AAV63753	Human JAG
44	155	25.1	5590	8	ACD06196	Human cDN
45	155	25.1	5896	7	ABZ34802	Coding se

ALIGNMENTS

RESULT 1
ABS52100
ID ABS52100 standard; DNA; 8354 BP.
XX
AC ABS52100;

DT 05-NOV-2002 (first entry)

XX Human TEN-M4-like gene.

XX Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis;
KW cell signal processing; metabolic pathway modulation; metabolic disorder;
KW obesity; diabetes; infectious disease; neurodegenerative disorder; acne;
KW Alzheimer's disease; Parkinson's disease; immune disorder; cancer;
KW haematopoietic disorder; cirrhosis; pancreatitis; learning defect;
KW memory defect; infertility; congenital heart defect; hair growth;
KW pigmentation disorder; endocrine disease; respiratory disease; health;
KW gastro-intestinal disease; reproductive; neurological disease;
KW bone marrow transplantation; endocrine disease; allergy; inflammation;
KW nephrological disorder; urinary system disorder; age-related disorder;
KW neuropsychiatric disorder; EGF-related protein; SCUBE1; TEN-M4;
KW adipocyte complement-related Clq tumour necrosis factor; out at first;
KW beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter;
KW type 1a membrane sushi-containing domain; butyrophilin;
KW single nucleotide polymorphism.

XX Homo sapiens.

XX Key Location/Qualifiers
FH variation replace(117,G)
FT /*tag= a

```
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(225,C)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism (SNP)"
FT replace(260,A)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism (SNP)"
XX
XX WO200257453-A2.
XX
XX 25-JUL-2002.
XX
XX 19-DEC-2001; 2001WO-US050331.
XX
XX 19-DEC-2000; 2000US-0256704P.
XX 20-DEC-2000; 2000US-0257314P.
XX 02-MAY-2001; 2001US-0288153P.
XX 29-MAY-2001; 2001US-0294075P.
XX 24-JUL-2001; 2001US-0307506P.
XX 10-AUG-2001; 2001US-0311590P.
XX 10-AUG-2001; 2001US-0311613P.
XX 29-AUG-2001; 2001US-0315617P.
XX 14-SEP-2001; 2001US-0322358P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
XX Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zerhusen BD, Liu X;
XX Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
XX
XX WPI; 2002-590744/63.
XX P-PSDB; ABG70388.
XX
XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
XX atherosclerosis, metabolic disorders, diabetes, obesity, infectious
XX disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
XX cancer.
XX
XX Claim 8; Page 50-52; 318pp; English.
XX
XX The present invention relates to new NOVX polypeptides. The invention is
XX useful for treating or preventing a NOVX-associated disorder such as
XX cardiomyopathy or atherosclerosis, where the disorder is related to cell
XX signal processing and metabolic pathway modulation in a subject,
XX preferably human. The invention is also useful for treating metabolic
XX disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
XX disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
XX disorders, haematopoietic disorders and various cancers. The molecules of
XX the invention are also useful for treating or preventing cirrhosis,
XX pancreatitis, learning and memory defects, infertility, congenital heart
XX defects, acne, hair growth, pigmentation disorders, endocrine disorders,
XX respiratory disease, gastro-intestinal diseases, reproductive, health,
XX neurological diseases, bone marrow transplantation, endocrine diseases,
XX allergy and inflammation, nephrological disorders, urinary system
XX disorders, neuropsychiatric disorders and age-related disorders. The
XX present nucleic acid sequence represents a NOVX gene. This sequence
XX encodes a NOVX protein of the invention
XX
XX SQ Sequence 8354 BP; 1968 A; 2437 C; 2271 G; 1678 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1.57e-40 Length: 8354
XX Score: 617.00 Matches: 101
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-10-029-020-14_COPY_750_850 (1-101) x ABS52100 (1-8354)
XX
XX Qy 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
XX |||||||
XX 2282 TGGATGGGGCAGCCTGCACGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACC 2341
```

```
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db |||||||
Db 2342 TGCCGCGACGGCAAGTGGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGCT 2401
QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db |||||||
Db 2402 CACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCTGGGTTGTCAATGGCAACGGCAGA 2461
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db |||||||
Db 2462 TGTACCTTAGACCTGAATGGTTGGCACTGCGCTGCGCTGGAGAGGAGCTGGC 2521
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db |||||||
Db 2522 TGTGACACTTCCATGGAGACTGCTGCGTGCACAGCAAAAGACAATGATGGAGATGGCCTG 2581
QY 101 Val 101
Db |||
Db 2582 GTG 2584
RESULT 2
ABS78652
ID ABS78652 standard; cDNA; 8645 BP.
XX
XX AC ABS78652;
XX
XX DT 16-DEC-2002 (first entry)
XX
XX DE Human cDNA encoding CGDD10, INCYTE 7488573CB1.
XX
XX KW Human; ss; gene; cell growth; differentiation; death; CGDD; cancer;
XX cell proliferative disorder; arteriosclerosis; atherosclerosis;
XX cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
XX polycythaemia vera; primary thrombocytopaenia; developmental disorder;
XX renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
XX neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
XX reproductive disorder; infertility; autoimmune disorder; gout; allergy;
XX inflammatory disorder; acquired immunodeficiency syndrome; infection;
XX autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
XX diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
XX multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
XX rheumatoid arthritis.
XX
XX OS Homo sapiens.
XX
XX XX WO200272830-A2.
XX
XX PD 19-SEP-2002.
XX
XX PF 08-FEB-2002; 2002WO-US003715.
XX
XX XX 09-FEB-2001; 2001US-0268111P.
XX 23-FEB-2001; 2001US-0271175P.
XX 08-MAR-2001; 2001US-0274503P.
XX 09-MAR-2001; 2001US-0274552P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX PI Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
XX Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
XX Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
XX
XX WPI; 2002-723356/78.
XX P-PSDB; ABG97359.
XX
XX New human proteins associated with cell growth, differentiation and
XX death, useful for diagnosing, treating or preventing autoimmune or
XX inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
XX atherosclerosis or hepatitis.
XX
XX Claim 5; Page 175-178; 181pp; English.
XX
```

CC The invention relates to an isolated polypeptide comprising CGDD1-12
CC (cell growth, differentiation and death), a naturally occurring amino
CC acid sequence at least 90% identical to CGDD, a biologically active
CC fragment or an immunogenic fragment. Also included are the
CC polynucleotides encoding CGDD1-12, a recombinant polynucleotide
CC comprising a promoter sequence operably linked to the CGDD
CC polynucleotides, a cell transformed with the recombinant polynucleotide,
CC a transgenic organism comprising the recombinant polynucleotide, an anti-
CC CGDD antibody, screening for compounds which bind to/modulate or are
CC ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a
CC CGDD polynucleotide microarray. The polypeptides, polynucleotides,
CC agonists and antagonists are useful for diagnosing, treating or
CC preventing disorders associated with aberrant expression of CGDD,
CC particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal haemoglobinuria, polycythaemia
CC vera, psoriasis, primary thrombocytopaenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anaemia or mental retardation),
CC neurological disorders (e.g. Alzheimer disease, Parkinson's disease or
CC epilepsy), reproductive disorders (e.g. infertility or a disruption in
CC the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS,
CC (acquired immunodeficiency syndrome) allergies, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis,
CC uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic
CC infections. They are also useful in the assessment of the effects of
CC exogenous compounds on the expression of nucleic acid and amino acid
CC sequences of proteins associated with CGDD. The present sequence encodes
CC a CGDD protein
XX
SQ Sequence 8645 BP; 2060 A; 2494 C; 2319 G; 1771 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 1.62e-40 Length: 8645
Score: 617.00 Matches: 101
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-029-020-14_COPY_750_850 (1-101) x ABS78652 (1-8645)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2352 TGGATGGGGGCAGCCTGCGACCAGCGGGCCTGCCACCCCGCGTGTGCCGAGCATGGGACC 2411
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2412 TGCCGCGACGGCAAGTGGAGTGCAGCCCTGGCTGGAAATGGCGAACACTGCACATCGCT 2471
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2472 CACTATCTGGATAGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAGA 2531
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2532 TGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGAGCTGGC 2591
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2592 TGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAGACAATGATGGAGATGGCCTG 2651
QY 101 Val 101
Db 2652 GTG 2654

RESULT 3
ABN85378
ID ABN85378 standard; DNA; 8438 BP.
XX
AC ABN85378;
XX
DT 21-OCT-2002 (first entry)

XX Human NOV1, TEN-M4 like protein, coding sequence.
DE
XX
KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
KW Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;
KW Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;
KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
KW Gene Therapy; NOV; cancer; heart disease; inflammation;
KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
KW asthma; IgA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
KW TEN-M4 like protein; chromosome 11; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 4..8395
FT /*tag= a
FT /trans_except= (pos: 1138..1147,aa:Met)
FT /product= "NOV1 protein"

WO200255704-A2.

18-JUL-2002.

09-JAN-2002; 2002WO-US0000554.

09-JAN-2001; 2001US-0260417P.

10-JAN-2001; 2001US-0260831P.

28-FEB-2001; 2001US-0272338P.

09-MAR-2001; 2001US-0274876P.

18-APR-2001; 2001US-0284704P.

(CURA-) CURAGEN CORP.

Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM;
Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
Macdougall J, Malyankar U, Millet I, Peyman J, Smithson G;
Gunther E, Stone DJ;

WPI; 2002-590674/63.

P-PSDB; ABB98401.

NOVX polypeptides and encoding polynucleotides, useful for preventing or
treating NOVX-associated disorders e.g. cancer, inflammation, or
Alzheimer's disease, and in chromosome mapping, tissue typing or
pharmacogenomics.

Claim 9; Page 8-9; 358pp; English.

The present sequence is a coding sequence for a NOV protein. The NOV
proteins and coding sequences are useful for treating or preventing NOV-
associated disorders or in the manufacture of a medicament for treating
the disorders, such as cancer, heart disease, inflammation, autoimmune
disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
IgA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
(e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
and other wasting disorders associated with chronic diseases. NOV1 is a
TEN-M4 like protein and the NOV1 gene is localised to chromosome 11

Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.97e-37 Length: 8438
Score: 579.00 Matches: 98
Percent Similarity: 94.29% Conservative: 1
Best Local Similarity: 93.33% Mismatches: 2
Query Match: 93.84% Indels: 4
DB: 6 Gaps: 2

US-10-029-020-14_COPY_750_850 (1-101) x ABN85378 (1-8438)

QY 75 GlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
Db 2322 GGATGGAGAGGAGCAGGCTGTGACGTAGCCATGGAGACTCTTTGCACAGATAGCAAGGAC 2381

QY 95 AsnAspGlyAspGlyLeuVal 101
Db 2382 AATGAAGGGATGGACTCATT 2402

RESULT 5
ABQ82344
ID ABQ82344 standard; cDNA; 8645 BP.

AC ABQ82344;

XX 17-DEC-2002 (first entry)

XX Human NOV15b encoding cDNA SEQ ID NO:37.

KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.

XX Homo sapiens.

XX
FH Key Location/Qualifiers
CDS 151..8316
FT /*tag= a
FT /product= "NOV15b"

XX WO200262999-A2.

XX 15-AUG-2002.

XX 31-DEC-2001; 2001WO-US049976.

XX 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.

XX (CURA-) CURAGEN CORP.

XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;

XX WPI; 2002-732706/79.
DR P-PSDB; ABP53587.

XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune

PT diseases.

PS Claim 8; Page 114-117; 444pp; English.

XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15b, which is
CC located on chromosome 4

XX SQ Sequence 8645 BP; 2343 A; 2060 C; 2200 G; 2038 T; 0 U; 4 Other;

Alignment Scores:

Pred. No.: 3.29e-32 Length: 8645
Score: 515.00 Matches: 84
Percent Similarity: 85.05% Conservative: 7
Best Local Similarity: 78.50% Mismatches: 10
Query Match: 83.47% Indels: 6
DB: 6 Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x ABQ82344 (1-8645)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2254 TGGACGGGCCAGCCTGTAATCAGAGAGCCTGCCACCCCCCGCTGTCCGAGACGGGACC 2313
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2314 TGCAAGGATGGCAAGTGTGAATGCAGCCAGGCGTGGAAATGGAGAGCACTGCACTATCGCT 2373
QY 41 HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
Db 2374 CACTATTTGGATAAGATAGTTAAAGACAAGATAGGATATAAAGAGGGTTGTCTGGTCTG 2433
QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 2434 TGCAACAGCAATGGAAGATGTACCTGGACCAAAATGGCGGACATTGTGTGTGCCAGCCT 2493
QY 75 GlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
Db 2494 GGATGGAGAGGAGCAGGCTGTGACGTAGCATGGAGACTCTTTGCACAGATAGCAAGGAC 2553
QY 95 AsnAspGlyAspGlyLeuVal 101
Db 2554 AATGAAGGGATGGACTCATT 2574
RESULT 6
ABQ82343
ID ABQ82343 standard; cDNA; 8675 BP.
XX
AC ABQ82343;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15a encoding cDNA SEQ ID NO:35.

XX Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disease; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 151..8328
FT FT /*tag= a
FT FT /product= "NOV15a"
FT FT /transl_except= (pos:1249..1251,aa:Ser)
XX
PN WO200262999-A2.
XX
PN 15-AUG-2002.
XX
PD 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53586.
XX
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 110-112; 444pp; English.
XX
XX The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's

CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15a, which is
CC located on chromosome 4
XX
SQ Sequence 8675 BP; 2362 A; 2067 C; 2201 G; 2041 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 3.3e-32 Length: 8675
Score: 515.00 Matches: 84
Percent Similarity: 85.05% Conservative: 7
Best Local Similarity: 78.50% Mismatches: 10
Query Match: 83.47% Indels: 6
DB: Gaps: 1

US-10-029-020-14_COPY_750_850 (1-101) x ABQ82343 (1-8675)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2266 TGGACGGGCCCGCCAGCCTGTAATCAGAGAGCCTGCCACCCCGCTGTGCCGAGCACGGGACC 2325

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2326 TGC AAGGATGGCAAGTGTGAATGCAGCCAGGCTGGAATGGAGAGCAGTGCCTATCGCT 2385

QY 41 HisTyrLeuAspArgValVal-----LysGluGlyCysProGlyLeu 54
Db 2386 CACTATTGGATAAGATAGTTAAAGACAAGATAGGATATAAAGAGGGTTGTCTGTCTG 2445

QY 55 CysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeu 74
Db 2446 TGC AACAGCAATGGAAGATGTACCTTGGACCAAAATGGCGGACATTGTGTGCCAGCCT 2505

QY 75 GlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSerLysAsp 94
Db 2506 GGATGGAGAGGAGCAGGCTGTGACGTAGCCATGGAGACTCTTGCACAGATAGCAAGGAC 2565

QY 95 AsnAspGlyAspGlyLeuVal 101
Db 2566 AATGAAGGGGATGGACTCATT 2586

RESULT 7
ABQ82346
ID ABQ82346 standard; cDNA; 8487 BP.
XX
AC ABQ82346;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15d encoding cDNA SEQ ID NO:41.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disease; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; gene; chromosome 4; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

FT CDS 299..8140
FT /*tag= a
FT /product= "NOV15d"
XX
PN WO200262999-A2.
XX PD
XX 15-AUG-2002.
XX
XX 31-DEC-2001; 2001WO-US049976.
PF PR 29-DEC-2000; 2000US-0258928P.
XX PR 02-JAN-2001; 2001US-0259415P.
PR PR 04-JAN-2001; 2001US-0259785P.
PR PR 20-FEB-2001; 2001US-0269814P.
PR PR 09-MAR-2001; 2001US-0279863P.
PR PR 29-MAR-2001; 2001US-0279832P.
PR PR 29-MAR-2001; 2001US-0279833P.
PR PR 13-APR-2001; 2001US-0283889P.
PR PR 18-APR-2001; 2001US-0284447P.
PR PR 25-APR-2001; 2001US-0286683P.
PR PR 29-MAY-2001; 2001US-0294080P.
PR PR 16-AUG-2001; 2001US-0312915P.
PR PR 17-AUG-2001; 2001US-0313325P.
PR PR 17-SEP-2001; 2001US-0322699P.
PR PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malyankar U, Shimkets RA, Tchernev VT, Spaderna SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.
DR P-PSDB; ABP53589.
XX
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
PT associated disorders, such as cancers, neurological disorders, disorders
PT of vesicular transport, gastrointestinal disorders, and autoimmune
PT diseases.
XX
PS Claim 8; Page 123-125; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cytosstatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence encodes human NOV15d, which is
CC located on chromosome 4
XX
SQ Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;

Alignment Scores: 3.03e-29 Length: 8487
Pred. No.: 478.50 Matches: 77
Score:

Percent Similarity: 81.19% Conservative: 5
Best Local Similarity: 76.24% Mismatches: 10
Query Match: 77.55% Indels: 9
DB: 6 Gaps: 1
US-10-029-020-14_COPY_750_850 (1-101) x ABQ82346 (1-8487)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2123 TGGACGGGCCCCAACCTGTAATCAGAGAGCCCTGCCACCCCGCTGTGCGGAGCACGGGACC 2182
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2183 TGCAAGGATGGCAAGTGTGAATGCAGCCATGGCTGGAATGGAGAGCACTGCACTATC--- 2239
QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2240 -----GAGGGTTGTCTCTGGTCTGTGCAACAGCAATGGAAGA 2275
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2276 TGTACCTGGACCAAAATGGCTGGCATTTGTGTGCCAGCCCTGGATGGAGAGGAGCAGGC 2335
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnGlyAspGlyLeu 100
Db 2336 TGTGACGTAGCCATGGAGACTCTTTGCACAGATAGCAAGGACAATGAAGGAGATGGACTC 2395
QY 101 Val 101
Db 2396 ATT 2398
RESULT 8
ABK92230
ID ABK92230 standard; DNA; 12879 BP.
XX
AC ABK92230;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated DNA sequence #116.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;
KW gene therapy; gene; ds.
XX Mammalia.
OS
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
DR WPI; 2002-471335/50.
DR P-PSDB; ABG61913.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a

PT prostate tissue.

XX

PS Claim 22; Page 394-397; 436pp; English.

XX

CC The present invention relates to methods of detecting a prostate cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with prostate cancer-associated polynucleotides (designated PC genes) that selectively hybridise to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g. mice, sheep and dogs). The methods of the invention are useful for diagnosing and treating prostate cancer in mammals. The prostate cancer-associated genes are useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications.

CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide sequences

XX

SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.46e-28	Length:	12879
Score:	464.50	Matches:	71
Percent Similarity:	82.00%	Conservative:	11
Best Local Similarity:	71.00%	Mismatches:	17
Query Match:	75.28%	Indels:	1
DB:	6	Gaps:	1

US-10-029-020-14_COPY_750_850 (1-101) x ABK92230 (1-12879)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||||
:::|
2210 TGGGTAGGACCAACATGTGAGGAACGCTCTCTCATTTCTGACTGAGCATGGCCAA 2269
|||||
:::|

Db

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||||
:::|
2270 TGCAAGATGGAAATGTAGTGTAGCCCTGGATGGGAGGGCGACCACTGCACAAATTGCT 2329
|||||
:::|

Db

QY 41 HistyLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||||
:::|
2330 CACTACTTAGAT---GCTGTCCGAGATGGCTGCCAGGGCTCTGCTTGGAAATGGACGA 2386
|||||
:::|

Db

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||||
:::|
2387 TGTACCTTGGATCAAAATGGTTGGCACTGTGTGTCTCAGGTGGGTGGAGTGGACAGGC 2446
|||||
:::|

Db

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
|||||
:::|
2447 TGCAATGTTGTCATGGAAATGCTTTTGGAGATAACTTGGACAATGATGGAGATGGTTTA 2506
|||||
:::|

Db

RESULT 9

AAK51828

ID AAK51828 standard; cDNA; 13202 BP.

XX

AC AAK51828;

XX

DT 06-NOV-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 373.

XX

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; ss.

XX

OS Homo sapiens.

XX

PN WO200157190-A2.

XX

PD 09-AUG-2001.

XX

PF 05-FEB-2001; 2001WO-US004098.

XX

PR 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.

PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.

PR 30-NOV-2000; 2000US-00728422.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX

DR WPI; 2001-476283/51.

DR P-PSDB; AAM78695.

XX

PT Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

PT

XX

PS Claim 1; Page 1414-1426; 6221pp; English.

XX

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

XX

SQ Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.63e-28	Length:	13202
Score:	464.50	Matches:	71
Percent Similarity:	82.00%	Conservative:	11
Best Local Similarity:	71.00%	Mismatches:	17
Query Match:	75.28%	Indels:	1
DB:	4	Gaps:	1

US-10-029-020-14_COPY_750_850 (1-101) x AAK51828 (1-13202)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
|||||
:::|
2536 TGGGTAGGACCAACATGTGAGGAACGCTCTCTCATTTCTGACTGAGCATGGCCAA 2595
|||||
:::|

Db

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
|||||
:::|
2596 TGCAAGATGGAAATGTGAGTGTAGCCCTGGATGGGAGGGCGACCACTGCACAAATTGCT 2655
|||||
:::|

Db

QY 41 HistyLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
|||||
:::|
2656 CACTACTTAGAT---GCTGTCCGAGATGGCTGCCAGGGCTCTGCTTGGAAATGGACGA 2712
|||||
:::|

Db

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
|||||
:::|
2713 TGTACCTTGGATCAAAATGGTTGGCACTGTGTGTCTCAGGTGGGTGGAGTGGACAGGC 2772
|||||
:::|

Db

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
|||||
:::|
2773 TGCAATGTTGTCATGGAAATGCTTTTGTGGAGATAACTTGGACAATGATGGAGATGGTTTA 2832
|||||
:::|

Db

RESULT 10

AAS14085
ID AAS14085 standard; DNA; 9826 BP.
XX
AC AAS14085;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human FCTR3b DNA sequence.
XX
KW Human; FCTRX; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
KW astrocytoma; congenital neonatal alloimmune thrombocytopaenia; infection;
KW neurological disorder; neurodegenerative disorders; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
KW mental health condition; immunological disorder; allergy; infertility;
KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
KW Corneal dystrophy-Greonouw type I; Corneal dystrophy-lattice type I;
KW Reis-Bucklers corneal dystrophy; cytostatic; immunosuppressive;
KW antiallergic; antiasthmatic; antiinfertility; antiinflammatory;
KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
KW gene therapy; FCTR3b; neurestin-like protein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..279
FT /*tag= b
FT CDS 280..8481
FT /*tag= a
FT /*product= "Human FCTR3b"
FT 3'UTR 8482..9826
FT /*tag= c
XX
PN WO200166747-A2.
XX
PN 13-SEP-2001.
XX
PF 05-MAR-2001; 2001WO-US0007160.
XX
PR 03-MAR-2000; 2000US-0186592P.
PR 03-MAR-2000; 2000US-0186718P.
PR 06-MAR-2000; 2000US-0187293P.
PR 06-MAR-2000; 2000US-0187294P.
PR 17-MAR-2000; 2000US-0190400P.
PR 07-APR-2000; 2000US-0196018P.
PR 03-JAN-2001; 2001US-0259548P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
XX
DR WPI; 2001-596837/67.
DR P-PSDB; AAU08680.
XX
PT Novel polypeptides designated as FCTRX polypeptides; useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
PS Claim 9; Page 33-35; 215pp; English.
XX
CC The invention relates to human FCTRX polypeptides, FCTR1-FCTR7, and the
CC nucleic acids encoding them. These sequences are useful for the treatment
CC or prevention of numerous disorders including myelogenous leukaemia,
CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
CC alloimmune thrombocytopaenia, neurological disorders, neurodegenerative
CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological

CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greonouw type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC FCTR3b, a neurestin-like protein
XX
SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3.49e-27 Length: 9826
Score: 454.00 Matches: 71
Percent Similarity: 77.23% Conservative: 7
Best Local Similarity: 70.30% Mismatches: 23
Query Match: 73.58% Indels: 0
DB: 5 Gaps: 0
US-10-029-020-14_COPY_750_850 (1-101) x AAS14085 (1-9826)
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2434 TGGACAGGCGCAGCGTGTGACCGAGCGGTGTGCCACCCCGCTGCATTGAGCACGGGACC 2493
QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2494 TGTAAAGATGGCAAATGTGAATCCGAGAGGGCTGGATGGTGAACACTGCACCATGGT 2553
QY 41 HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2554 AGGCAAACGGCAGGCCACCGAAACAGATGGCTGCCCTGACTTGTCAACGGTAACGGGAGA 2613
QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2614 TGCACACTGGGTTCAGAACAGCTGGCAGTGTCTGCCAGACCGGCTGGAGAGGGCCCGGA 2673
QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2674 TGC AACGTTGCCATGGAAACTTCCTGTGTGATACAAAGGATAATGAGGGAGATGGCCTG 2733
QY 101 Val 101
Db 2734 GTG 2736
RESULT 11
ADB32023
ID ADB32023 standard; cDNA; 9826 BP.
XX
AC ADB32023;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human FCTR3b cDNA.
XX
KW Human; gene; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
XX
OS Homo sapiens.
XX

XX Novel polypeptides designated as FCTR_X polypeptides, useful in detection,
PT prevention and treatment of a broad range of pathological states.
XX
XX
PS Claim 9; Page 37-39; 215pp; English.
XX
CC The invention relates to human FCTR_X polypeptides, FCTR1-FCTR7, and the
CC nucleic acids encoding them. These sequences are useful for the treatment
CC or prevention of numerous disorders including myelogenous leukaemia,
CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
CC alloimmune thrombocytopaenia, neurological disorders, neurodegenerative
CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
CC myelodysplastic syndrome, mental health conditions, immunological
CC disorders, allergy and infection, bronchial asthma, Avellino type
CC eosinophilia, lung diseases, reproductive disorders, infertility, male
CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
CC infection Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
CC Corneal dystrophy-Greenouw type I, Corneal dystrophy-lattice type I and
CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
CC the FCTR3a homologue FCTR3f
XX
SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.7e-27	Length:	9729
Score:	448.50	Matches:	71
Percent Similarity:	77.23%	Conservative:	7
Best Local Similarity:	70.30%	Mismatches:	14
Query Match:	72.69%	Indels:	9
DB:	5	Gaps:	1

US-10-029-020-14_COPY_750_850 (1-101) x AAS14089 (1-9729)

QY	1	TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr	20
Db	2364	TGGACAGGCGCAGCGTGTGACCGCGTGTGCCACCCCGCTGCATTGAGCATGGGACC	2423
QY	21	CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla	40
Db	2424	TGTAAGATGGCAATGTGAATGCCGAGAGGGCTGGATGGTGAACACTGCACATT---	2480
QY	41	HisTyrLeuAspArgValValLysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg	60
Db	2481	-----GATGGCTGCCCTGACTTGTGCAACGGTAACGGGAGA	2516
QY	61	CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly	80
Db	2517	TGCACACTGGGTGAGAACAGCTGGCAGTGTCTGCCAGACCGGCTGGAGAGGCCCGGA	2576
QY	81	CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu	100
Db	2577	TGCAACGTTGCCATGGAAACTTCCTGTCTGATAACAAGGATAATGAGGGAGATGGCCTG	2636
QY	101	Val 101	
Db	2637	GTG 2639	

ADB32028 standard; cDNA; 9729 BP.

ADB32028;

04-DEC-2003 (first entry)

Human FCTR3f cDNA.

Human; gene; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;

renal cell carcinoma; melanoma; clear cell carcinoma;
granular cell carcinoma; neurological disorder;
neurodegenerative disorder; nerve trauma;
familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
Gardner syndrome; mental health condition; immunological disorder;
allergy; asthma; lung disease; reproductive disorder; deafness;
glycoprotein deficiency; desmoid tumour; turcot syndrome;
liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
plasmodium falciparum infection; Greenouw's corneal dystrophy;
lattice corneal dystrophy.
Homo sapiens.
US2003087816-A1.
08-MAY-2003.
05-MAR-2001; 2001US-00800198.
03-MAR-2000; 2000US-0186592P.

(VERM/)	VERMET C.
(FERN/)	FERNANDES E.
(SHIM/)	SHIMKETS R.
(HERR/)	HERRMANN J.
(MAJU/)	MAJUMDER K.
(MACD/)	MACDOUGALL J.
(MISH/)	MISHRA V.
(MEZE/)	MEZES P S.
(RAST/)	RASTELLI L.

Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K;
Macdougall J, Mishra V, Mezes PS, Rastelli L;

WPI; 2003-625633/59.
P-PSDB; ADB32029.

New FCTR_X polypeptide and encoding polynucleotide, useful for preventing
or treating FCTR_X-related disorders, such as cancer, autoimmune,
neurodegenerative, gastrointestinal, reproductive and inflammatory
diseases.

Claim 9; Page 34-37; 155pp; English.

The invention relates to FCTR_X polypeptides and the polynucleotides
encoding them. The sequences of the invention are useful for the
manufacture of a medicament for diagnosing and treating disorders
associated with the FCTR_X polypeptide, such as colorectal cancer,
adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
tumours, mammary tumours, human gliomas, astrocytomas, renal cell
carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
cell and granular cell carcinomas, neurological disorders.
neurodegenerative disorders, nerve trauma, familial myelodysplastic
syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
conditions, immunological disorders, allergy and infection, asthma, lung
diseases, male and female reproductive disorders, deafness, glycoprotein
deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis
C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
infection, spinocerebellar ataxia, Plasmodium falciparum infection,
Greenouw's corneal dystrophy and lattice corneal dystrophy. This sequence
represents cDNA encoding an FCTR_X polypeptide of the invention.

Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	9.7e-27	Length:	9729
Score:	448.50	Matches:	71
Percent Similarity:	77.23%	Conservative:	7
Best Local Similarity:	70.30%	Mismatches:	14
Query Match:	72.69%	Indels:	9
DB:	9	Gaps:	1

US-10-029-020-14_COPY_750_850 (1-101) x ADB32028 (1-9729)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCysHisProArgCysAlaGluHisGlyThr 20
Db 2364 TGGACAGGCGCAGCGTGTGACCGCGTGTGCCACCCCGCGTGCATTGAGCATGGGACC 2423

QY 21 CysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIleAla 40
Db 2424 TGTAAAGATGGCAATGTGAATGCCGAGAGGGCTGGAATGGTGAACACTGCACCATTT--- 2480

QY 41 HisTyrLeuAspArgValVallysGluGlyCysProGlyLeuCysAsnGlyAsnGlyArg 60
Db 2481 -----GATGGCTGCCCTGACTTGTGCAACGGTAACGGGAGA 2516

QY 61 CysThrLeuAspLeuAsnGlyTrpHisCysValCysGlnLeuGlyTrpArgGlyAlaGly 80
Db 2517 TGCACACTGGGTGCAGAACAGCTGGCAGTGTGTCTGCCAGACCGGCTGGAGAGGGCCCGGA 2576

QY 81 CysAspThrSerMetGluThrAlaCysGlyAspSerLysAspAsnAspGlyAspGlyLeu 100
Db 2577 TGCAACGTTGCCATGGAACTTCCTGTCTGTATACAAGGATAATGAGGGAGATGGCCTG 2636

QY 101 Val 101
Db 2637 GTG 2639

RESULT 14
AAL60066
ID AAL60066 standard; cDNA; 4245 BP.
XX
AC AAL60066;
XX
DT 27-AUG-2003 (first entry)
XX
DE Human Pc099 cDNA.
XX
KW Human; differentially regulated protein; prevention; therapy; vaccine;
KW prostate cancer; gene therapy; Pc099; teneurin-2; chromosome 5q34; gene;
KW ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..4056
FT /*tag= a
FT /product= "Human Pc099 protein"
FT /note= "No start codon"
FT /partial
XX
PN WO2003040331-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-US035563.
XX
PR 07-NOV-2001; 2001US-0331041P.
PR 07-NOV-2001; 2001US-0331042P.
PR 18-DEC-2001; 2001US-0340251P.
PR 07-JAN-2002; 2002US-0344791P.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Jay G, Kovacs KF, Fan W;
XX
DR WPI; 2003-449451/42.
DR P-PSDB; AAO29571.
XX
PT New polynucleotide for diagnosing, staging, monitoring, prognosticating,
PT preventing or treating, or determining the predisposition to, diseases or
PT conditions such as prostate cancer, and for research or forensic science.
XX
PS Claim 29; Page 158-163; 100pp; English.

XX The present invention relates to novel differentially regulated genes and
CC polypeptides encoded by them. Sequences of the invention are useful in
CC diagnosing, staging, monitoring, prognosticating, preventing, treating or
CC determining the predisposition to diseases or conditions such as prostate
CC cancer. They may be used as molecular markers, drug targets, vaccines, in
CC gene therapy, research, clinical medicine or forensic science. The
CC present sequence is a differentially regulated prostate cDNA, Pc099 which
CC codes for teneurin-2. Pc099 gene is located on chromosome 5q34

XX
SQ Sequence 4245 BP; 1035 A; 1132 C; 1115 G; 963 T; 0 U; 0 Other;

Alignment Scores:
Pred. NO.: 1.66e-17 Length: 4245
Score: 330.50 Matches: 60
Percent Similarity: 57.36% Conservative: 14
Best Local Similarity: 46.51% Mismatches: 26
Query Match: 53.57% Indels: 29
DB: 7 Gaps: 4

US-10-029-020-14_COPY_750_850 (1-101) x AAL60066 (1-4245)

QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCys---HisProArgCysAlaGluHisGly 19
Db 790 TACAAAGCGGAGCACTGTGAGGAAGTTGATTGCTTGGATCCACCTGTCTCCAGCCAGGA 849

QY 20 ThrCysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle 39
Db 850 GTCTGTGTGAATGGAGATGCCTGTGCAGCCCTGGCTGGGTGGTCTGAACCTGTGAGCTG 909

QY 40 AlaHis-----TyrLeu----- 43
Db 910 GCAGGGTCCAGTGCCCGACAGCAGTGCGAGTGGCATGGCAGTACCTGCCTGACACGGGC 969

QY 44 -----AspArgValVallysGluGlyCysPro 52
Db 970 CTCTGCAGCTGCGATCCCAACTGGATGGTCCCGACTGCTCTGTT---GATGGCTGCCCT 1026

QY 53 GlyLeuCysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCys 72
Db 1027 GACTTGTGCAACGGTAACGGGAGATGCACACTGGGTGAGTGGTGCAGACAGCTGGCAGTGTCTGTC 1086

QY 73 GlnLeuGlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSer 92
Db 1087 CAGACCGGCTGGAGAGGGCCCGGATGCAACGTTGCCATGGAACTTCCTGTCTGTGATAAC 1146

QY 93 LysAspAsnAspGlyAspGlyLeuVal 101
Db 1147 AAGGATAATGAGGGAGATGGCCTGGTG 1173

RESULT 15
ACC72051
ID ACC72051 standard; DNA; 9058 BP.
XX
AC ACC72051;
XX
DT 08-JUL-2003 (first entry)
XX
DE BCU0205A gene #SEQ ID 79.
XX
KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
KW drug discovery; clinical medicine; forensic medicine; gene;
KW chromosome 5q33.3; ds.
XX
OS Homo sapiens.
XX
PN WO2003029421-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326526P.


```
PR 14-MAY-2002; 2002US-00144194.
XX
PA (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
DR WPI; 2003-381623/36.
DR P-PSDB; ABR58317.
XX
PT New isolated human differentially-regulated breast cancer polynucleotide
PT and polypeptide, useful for diagnosing, staging, prognosticating,
PT preventing and/or treating diseases and conditions relating to breast
PT cancer.
XX
PS Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
CC The invention relates to isolated polynucleotides which are
CC differentially-regulated in breast cancer. The methods and compositions
CC of the present invention are useful for detecting, diagnosing, staging,
CC monitoring, prognosticating, preventing and/or treating diseases and
CC conditions relating to breast cancer, and may be used in gene therapy or
CC antisense therapy. They can also be used in research, drug discovery,
CC clinical medicine and forensic medicine. Sequences given in records
CC ACC72012-ACC72074 represent polynucleotides of the invention that are
CC differentially-regulated in breast cancer. NOTE: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.:	3.67e-17	Length:	9058
Score:	330.50	Matches:	60
Percent Similarity:	57.36%	Conservative:	14
Best Local Similarity:	46.51%	Mismatches:	26
Query Match:	53.57%	Indels:	29
DB:	7	Gaps:	4

US-10-029-020-14_COPY_750_850 (1-101) x ACC72051 (1-9058)

```
QY 1 TrpMetGlyAlaAlaCysAspGlnArgAlaCys---HisProArgCysAlaGluHisGly 19
Db 1571 TACAAAGGCGGAGCACTGTGAGGAAGTTGATTGCTTGATCCACCTGCTCCAGCCACGGA 1630
QY 20 ThrCysArgAspGlyLysCysGluCysSerProGlyTrpAsnGlyGluHisCysThrIle 39
Db 1631 GTCTGTGTGAATGGAGATGCCTGTGCAGCCCTGGCTGGGGTGGTCTGAACTGTGAGCTG 1690
QY 40 AlaHis-----TyrLeu----- 43
Db 1691 GCGAGGGTCCAGTGCCAGACCAGTGCAGTGGGCATGGCACGTACCTGCCTGACACGGGC 1750
QY 44 -----AspArgValValLysGluGlyCysPro 52
Db 1751 CTCTGCAGTGCATCCCAACTGGATGGGTCCCGACTGCTCTGTT---GATGCTGCCCT 1807
QY 53 GlyLeuCysAsnGlyAsnGlyArgCysThrLeuAspLeuAsnGlyTrpHisCysValCys 72
Db 1808 GAC TTGTGCAACGGTAACGGGAGATGCACACTGGGT CAGAACAGCTGGCAGTGTGTCTGC 1867
QY 73 GlnLeuGlyTrpArgGlyAlaGlyCysAspThrSerMetGluThrAlaCysGlyAspSer 92
Db 1868 CAGACCGGCTGGAGAGGGCCCGGATGCAACGTTGCCATGGAACCTTCCTGTGCTGATAAC 1927
QY 93 LysAspAsnAspGlyAspGlyLeuVal 101
Db 1928 AAGGATAATGAGGAGATGGCCTGGTG 1954
```

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:33:47 ; Search time 142.298 Seconds
(without alignments)
2448.158 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520

Perfect score: 372

Sequence: 1 TFRSQVFIDHPVHLKFNV.....RLLTQEARSLGTPRQSRGT 71

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3225727 seqs, 2453303834 residues

Total number of hits satisfying chosen parameters: 6451454

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112218_29331/app_query.fasta_1.3519
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10029020@cgn_1_1_2156@runat_06082004_112218_29331
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

ALIGNMENTS

RESULT 1
US-10-383-201-45
; Sequence 45, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957

1	372	100.0	1392	13	US-10-383-201-45	Sequence 45, Appl
2	372	100.0	1392	13	US-10-383-201-53	Sequence 53, Appl
3	372	100.0	1476	13	US-10-383-201-41	Sequence 41, Appl
4	372	100.0	1680	15	US-10-290-578-3	Sequence 3, Appl
5	372	100.0	2387	15	US-10-290-578-11	Sequence 11, Appl
6	372	100.0	3111	9	US-09-773-517-12	Sequence 12, Appl
7	372	100.0	3111	9	US-09-792-025-12	Sequence 12, Appl
8	372	100.0	3111	9	US-09-849-868-12	Sequence 12, Appl
9	372	100.0	3111	15	US-10-290-578-1	Sequence 1, Appl
10	372	100.0	3111	15	US-10-453-183-12	Sequence 12, Appl
11	372	100.0	8354	13	US-10-383-201-43	Sequence 43, Appl
12	372	100.0	8354	13	US-10-029-020-13	Sequence 13, Appl
13	372	100.0	8355	13	US-10-383-201-55	Sequence 55, Appl
14	372	100.0	8438	13	US-10-042-865-1	Sequence 1, Appl
15	254	68.3	8473	17	US-10-038-854-39	Sequence 39, Appl
16	254	68.3	8487	17	US-10-038-854-41	Sequence 41, Appl
17	254	68.3	8645	17	US-10-038-854-37	Sequence 37, Appl
18	234	62.9	8675	17	US-10-038-854-35	Sequence 35, Appl
19	220	59.1	12880	16	US-10-295-027-927	Sequence 927, App
20	194.5	52.3	1430	9	US-09-808-602-5	Sequence 5, Appl
21	194.5	52.3	1430	10	US-09-800-198-5	Sequence 5, Appl
22	194.5	52.3	1431	10	US-09-977-418-21	Sequence 21, Appl
23	194.5	52.3	1431	10	US-09-977-033A-21	Sequence 21, Appl
24	194.5	52.3	1431	10	US-09-977-751C-21	Sequence 21, Appl
25	194.5	52.3	1431	10	US-09-977-639A-21	Sequence 21, Appl
26	194.5	52.3	1431	11	US-09-977-819B-21	Sequence 21, Appl
27	194.5	52.3	1727	16	US-10-094-749-710	Sequence 710, App
28	194.5	52.3	1743	15	US-10-098-871-7	Sequence 7, Appl
29	194.5	52.3	2496	9	US-09-808-602-75	Sequence 75, Appl
30	194.5	52.3	2496	10	US-09-800-198-63	Sequence 63, Appl
31	194.5	52.3	8409	9	US-09-808-602-79	Sequence 79, Appl
32	194.5	52.3	8409	10	US-09-800-198-67	Sequence 67, Appl
33	194.5	52.3	8575	13	US-10-072-012-143	Sequence 143, App
34	194.5	52.3	8689	9	US-09-808-602-78	Sequence 78, Appl
35	194.5	52.3	8689	10	US-09-800-198-66	Sequence 66, Appl
36	194.5	52.3	8797	9	US-09-808-602-74	Sequence 74, Appl
37	194.5	52.3	8797	9	US-09-808-602-77	Sequence 77, Appl
38	194.5	52.3	8797	10	US-09-800-198-62	Sequence 62, Appl
39	194.5	52.3	8797	10	US-09-800-198-65	Sequence 65, Appl
40	194.5	52.3	9058	16	US-10-144-194A-79	Sequence 79, Appl
41	194.5	52.3	9695	16	US-10-144-194A-81	Sequence 81, Appl
42	194.5	52.3	9729	9	US-09-808-602-12	Sequence 12, Appl
43	194.5	52.3	9729	10	US-09-800-198-12	Sequence 12, Appl
44	194.5	52.3	9826	9	US-09-808-602-7	Sequence 7, Appl
45	194.5	52.3	9826	10	US-09-800-198-7	Sequence 7, Appl


```
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 45
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1392)
US-10-383-201-45

Alignment Scores:
Pred. No.:      1.01e-47      Length:      1392
Score:          372.00      Matches:      71
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:             13      Gaps:
US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-45 (1-1392)

Qy      1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db      244 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 303

Qy      21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db      304 CTGGAAAGGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCTCCCTCCTTCACATACA 363

Qy      41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db      364 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 423

Qy      61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db      424 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 456

RESULT 2
US-10-383-201-53
; Sequence 53, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 41
; LENGTH: 1476
; TYPE: DNA
```

```
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 53
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1386)
US-10-383-201-53

Alignment Scores:
Pred. No.:      1.01e-47      Length:      1392
Score:          372.00      Matches:      71
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:             13      Gaps:
US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-53 (1-1392)

Qy      1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db      244 ACTTCTGGAGATCTCAAGTGTTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 303

Qy      21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db      304 CTGGAAAGGCAGCCCTGGTTGGCATTATGGCAGAAAAGGCTCCCTCCTTCACATACA 363

Qy      41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db      364 CAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 423

Qy      61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db      424 CTAGAGGGGACCCCGCCAGTCTCGGGGAACT 456

RESULT 3
US-10-383-201-41
; Sequence 41, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; CURRENT FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 41
; LENGTH: 1476
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1464)
US-10-383-201-41
Alignment Scores:
Pred. No.: 11e-47 Length: 1476
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-41 (1-1476)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 322 ACTTTCGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 381
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 382 CTGGGAAAGGCAGCCCTGTTGGCATTATGGCAGAAAGGCCTCCCTCCCTTCACATACA 441
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 442 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGGCGGGAGC 501
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 502 CTAGAGGGGACCCCGCCGAGTCTCGGGGAACT 534
```

RESULT 4

```

US-10-290-578-3
; Sequence 3, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; Sliwkowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-No. US20030078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1680 base pairs
```

```

; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-290-578-3
Alignment Scores:
Pred. No.: 1.3e-47 Length: 1680
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-290-578-3 (1-1680)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTTCGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1407
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1408 CTGGGAAAGGCAGCCCTGTTGGCATTATGGCAGAAAGGCCTCCCTCCCTTCACATACA 1467
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgGlyLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGGCGGGAGC 1527
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1528 CTAGAGGGGACCCCGCCGAGTCTCGGGGAACT 1560
```

RESULT 5

```

US-10-290-578-11
; Sequence 11, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; Sliwkowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-No. US20030078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
```

TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-290-578-11

Alignment Scores:
Pred. No.: 2.08e-47 Length: 2387
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-290-578-11 (1-2387)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 853 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 912
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 913 CTGGAAAGGCAGCCCTGTTGGCAATTATGGCAGAAAGGCTCCCTCCCTTCACATACA 972
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 973 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1032
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1033 CTAGAGGGGACCCCGCCAGTCCTCGGGAACT 1065

RESULT 6
US-09-773-517-12
; Sequence 12, Application US/09773517
; Patent No. US20010023241A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/09/773,517
; CURRENT FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/243,198
; PRIOR FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-773-517-12

Alignment Scores:
Pred. No.: 2.96e-47 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-773-517-12 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 1740
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGAAAGGCAGCCCTGTTGGCAATTATGGCAGAAAGGCTCCCTCCCTTCACATACA 1800
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCCAGTCCTCGGGAACT 1893
RESULT 7
US-09-792-025-12
; Sequence 12, Application US/09792025
; Patent No. US20020042087A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/792,025
; FILING DATE: 23-Feb-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/020,598
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: P1145
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: nucleic acid
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-792-025-12

Alignment Scores:
Pred. No.: 2.96e-47 Length: 3111
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-792-025-12 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 1740
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGAAAGGCAGCCCTGTTGGCAATTATGGCAGAAAGGCTCCCTCCCTTCACATACA 1800
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1860
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71


```
Db      1861 CTAGAGGGACCCCGCCAGTCTCGGGAACT 1893
RESULT 8
US-09-849-868-12
; Sequence 12, Application US/09849868
; Patent No. US2002008129A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Gao, Wei-Qiang
; TITLE OF INVENTION: HAIR CELL DISORDERS
; FILE REFERENCE: GENENT.035VPC
; CURRENT APPLICATION NUMBER: US/09/849,868
; CURRENT FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)...(2637)
US-09-849-868-12

Alignment Scores:
Pred. No.:      2.96e-47      Length:      3111
Score:          372.00      Matches:      71
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              9          Gaps:          0

US-10-029-020-14_COPY_450_520 (1-71) x US-09-849-868-12 (1-3111)

QY      1 ThrPheTIpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db      1681 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1740

QY      21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db      1741 CTGGGAAAGGCAGCCCTGGTGGCATTATGGCAGAAAAGGCCTCCCTCCTTCACATACA 1800

QY      41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db      1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTAAACCAGGAGGCGCGGAGC 1860

QY      61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db      1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 9
US-10-290-578-1
; Sequence 1, Application US/10290578
; Publication No. US20030078389A1
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/10/290,578
; FILING DATE: 08-No. US20030078389A1-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 60/021640
; FILING DATE: 12-Jul-96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-290-578-1

Alignment Scores:
Pred. No.:      2.96e-47      Length:      3111
Score:          372.00      Matches:      71
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              15         Gaps:          0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-290-578-1 (1-3111)

QY      1 ThrPheTIpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db      1681 ACTTTCGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1740

QY      21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db      1741 CTGGGAAAGGCAGCCCTGGTGGCATTATGGCAGAAAAGGCCTCCCTCCTTCACATACA 1800

QY      41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db      1801 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTAACCAGGAGGCGCGGAGC 1860

QY      61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db      1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 10
US-10-453-183-12
; Sequence 12, Application US/10453183
; Publication No. US20030199429A1
; GENERAL INFORMATION:
; APPLICANT: Sliwowski, Mark
; APPLICANT: Kern, Jeffrey
; TITLE OF INVENTION: Use of Heregulin as a Growth Factor
; FILE REFERENCE: P1145R1
; CURRENT APPLICATION NUMBER: US/10/453,183
; CURRENT FILING DATE: 2003-06-03
; PRIOR APPLICATION NUMBER: US/09/243,198
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: US 60/073,866
; PRIOR FILING DATE: 1998-02-04
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 12
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-453-183-12
```

Alignment Scores: 2.96e-47 Length: 3111
Pred. No.: 372.00 Matches: 71
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 15

US-10-029-020-14_COPY_450_520 (1-71) x US-10-453-183-12 (1-3111)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1681 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 1740

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1741 CTGGAAAGGCAGCCCTGGTGGCATTATGGCAGAAAAGGCCTCCCTCCTTCACATACA 1800

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1801 CAGTTTGACTTTGTGGAGTGTGGATGGCAGGAGGCTCCTAAACCCAGGAGCGCGGAGC 1860

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1861 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1893

RESULT 11
US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; PRIOR FILING DATE: 2003-03-06
; CURRENT FILING DATE: 2003-03-06
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (35)..(8341)
US-10-383-201-43

Alignment Scores: 1.11e-46 Length: 8354
Pred. No.: 372.00 Matches: 71
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00%

DB: 13 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-43 (1-8354)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1382 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 1441

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1442 CTGGAAAGGCAGCCCTGGTGGCATTATGGCAGAAAAGGCCTCCCTCCTTCACATACA 1501

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
Db 1502 CAGTTTGACTTTGTGGAGTGTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1561

QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1562 CTAGAGGGGACCCCGCCAGTCTCGGGAACT 1594

RESULT 12
US-10-029-020-13
; Sequence 13, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; PRIOR FILING DATE: 2001-12-19
; CURRENT FILING DATE: 2001-12-19
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Alignment Scores: 1.11e-46 Length: 8354
Pred. No.: 372.00 Matches: 71
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 13

US-10-029-020-14_COPY_450_520 (1-71) x US-10-029-020-13 (1-8354)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1382 ACTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTCT 1441

QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1442 CTGGAAAGGCAGCCCTGGTGGCATTATGGCAGAAAAGGCCTCCCTCCTTCACATACA 1501

QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1502 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1561
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71
Db 1562 CTAGAGGGGACCCCGCGCCAGTCTCGGGGAACT 1594
RESULT 13
US-10-383-201-55
; Sequence 55, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 8355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(8325)
US-10-383-201-55
Alignment Scores:
Pred. No.: 1,11e-46 Length: 8355
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-383-201-55 (1-8355)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1348 ACTTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1407
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
Db 1408 CTGGGAAGGCAGCCCTGGTTGGCATTATGGCAGAAAGGCCCTCCCTCTCACATACA 1467
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer 60
Db 1468 CAGTTTGACTTTGTGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGCGCGGAGC 1527
QY 61 LeuGluGlyThrProArgGlnSerArgGlyThr 71

Db 1528 CTAGAGGGGACCCCGCGCCAGTCTCGGGGAACT 1560
RESULT 14
US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; TITLE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1
Alignment Scores:
Pred. No.: 1.12e-46 Length: 8438
Score: 372.00 Matches: 71
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x US-10-042-865-1 (1-8438)
QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
Db 1358 ACTTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCATGTGTCT 1417


```
QY      21  LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
      |||
Db      1418 CTGGGAAGGCAGCCCTGGTTGGCATTATATGGCAGAAAAGGCCTCCCTCCCTCACATACA 1477

QY      41  GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSer 60
      |||
Db      1478 CAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCTAACCAGGAGGCGCGGAGC 1537

QY      61  LeuGluGlyThrProArgGlnSerArgGlyThr 71
      |||
Db      1538 CTAGAGGGGACCCCGCGCAGTCTCGGGGAACT 1570

RESULT 15
US-10-038-854-39
; Sequence 39, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 832.674 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_450_520
Perfect score: 372
Sequence: 1 TFWRSQVFIDHPVHLKENVS.....RLLTQEARSLGTPRQSRGT 71

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10029020@CGN_1_13135@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	302	81.2	870	10	BG036207	BG036207 602326960
2	248.5	66.8	2627	11	AK050784	AK050784 Mus muscu
3	228	61.3	701	28	AZ661742	AZ661742 1M0540B02
4	220.5	59.3	3447	11	AK047423	AK047423 Mus muscu
5	202.5	54.4	594	13	BQ257598	BQ257598 NISC_kp04
6	201	54.0	519	13	BX494675	BX494675 DKFZp779B
7	196.5	52.8	306	14	CF538752	CF538752 UI-M-G10-
8	194.5	52.3	2716	11	AK031198	AK031198 Mus muscu
9	186	50.0	1013	12	BM017187	BM017187 603643790
10	171	46.0	247	9	AL710528	AL710528 DKFZp686K
11	143	38.4	768	29	BX172779	BX172779 Danio rer
12	143	38.4	769	29	BX245366	BX245366 Danio rer
13	142	38.2	392	14	N67072	N67072 yz45c06.s1
14	139	37.4	749	29	BX178536	BX178536 Danio rer
15	117	31.5	340	28	CC122834	CC122834 NDL.45P17
16	83	22.3	1052	9	AA051165	AA051165 mj42a04.x
17	78.5	21.1	456	9	AI762835	AI762835 w104g12.x
18	75	20.2	562	29	DR20L14T	AL747302 Danio rer
19	74.5	20.0	547	10	BF498290	BF498290 AT12676.5
20	73	19.6	600	28	BZ894471	BZ894471 Hg13_0121
21	73	19.6	1148	12	BM801011	BM801011 AGENCOURT
22	72	19.4	644	28	AZ083567	AZ083567 RPCI-23-4
23	71	19.1	426	14	CB274460	CB274460 mai69d03.
24	71	19.1	433	29	BX179117	BX179117 Danio rer
25	71	19.1	583	12	BM355690	BM355690 rr28cl1.y
26	70.5	19.0	498	10	BF501431	BF501431 AT16761.5
27	70.5	19.0	1073	28	BZ559588	BZ559588 pacs2-164
28	70	18.8	348	13	BY046474	BY046474 BY046474
29	70	18.8	446	14	CD453939	CD453939 WHE0922.F
30	70	18.8	511	14	CB447588	CB447588 701597_MA
31	70	18.8	592	13	BQ777099	BQ777099 il43d04.y
32	70	18.8	720	28	AQ005704	AQ005704 CIT-HSP-2
33	69.5	18.7	1372	13	BU901558	BU901558 AGENCOURT
34	69	18.5	397	14	N77794	N77794 yz83g05.x1
35	69	18.5	414	9	AI173157	AI173157 vz78d11.x
36	69	18.5	438	12	BM841907	BM841907 K-EST0119
37	69	18.5	481	13	CA128790	CA128790 SCQGLR201
38	69	18.5	579	14	CD894862	CD894862 G118.127F
39	68.5	18.4	377	10	BF423355	BF423355 Hc d11.28
40	68.5	18.4	542	12	BJ203521	BJ203521 BJ203521
41	68	18.3	369	13	BY294778	BY294778 BY294778
42	68	18.3	476	13	BY031967	BY031967 BY031967
43	68	18.3	503	9	AJ434765	AJ434765 AJ434765
44	68	18.3	604	12	BJ606630	BJ606630 BJ606630
45	68	18.3	666	12	BI959265	BI959265 HVSMEN001

ALIGNMENTS

RESULT 1
BG036207
LOCUS
DEFINITION BG036207 870 bp mRNA linear EST 24-JAN-2001
602326960F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428351 5',
mRNA sequence.
ACCESSION BG036207
VERSION BG036207.1 GI:12431132
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)

misc_feature
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="9 days embryo"
1. .2627
/note="odd Oz/ten-m homolog 3 (Drosophila)
(MGJ|MGJ:1345183, GB|NM_011857, evidence: BLASTN, 99%,
match=1753)"

ORIGIN

Alignment Scores:
Pred. No.: 2.7e-22 Length: 2627
Score: 248.50 Matches: 48
Percent Similarity: 80.28% Conservative: 9
Best Local Similarity: 67.61% Mismatches: 11
Query Match: 66.80% Indels: 3
DB: 11 Gaps: 1

US-10-029-020-14_COPY_450_520 (1-71) x AK050784 (1-2627)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
|||||
Db 777 TTCTGGAGATCGCAGCTCTTTATTGATCAGCCACAGTTTCTTAAGTTCAACATCTCTTT 836
QY 22 GlyLysAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
|||||
Db 837 CAGAAGGATGATTCGAGTGTACGGCGGAGGGCTTACCGCTTCCCATACTCAG 896
QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuThrGlnGluAlaArgSerLeu 61
:::|||||
Db 897 TACGACTTGTGGAACTACTGGATGTAGCAGGTTAATTGCGAGAGAGCGGAACCTG 956
QY 62 -----GluGlyThrProArgGlnSerArg 69
|||
Db 957 GTGGAGTCCGAAGAGCGCGGCGGCGGAGGAG 989

RESULT 3
AZ661742
LOCUS
DEFINITION
1M0540B02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0540B02 R, genomic survey sequence.

ACCESSION
AZ661742
VERSION
GSS.
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 701)

REFERENCE
AUTHORS
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.

TITLE
Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL
COMMENT
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177

FEATURES
source
1. .701
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"

/clone="UUGC1M0540B02"
/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

ORIGIN

Alignment Scores:
Pred. No.: 2.39e-20 Length: 701
Score: 228.00 Matches: 45
Percent Similarity: 83.93% Conservative: 2
Best Local Similarity: 80.36% Mismatches: 9
Query Match: 61.29% Indels: 0
DB: 28 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x AZ661742 (1-701)

QY 1 ThrPheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSer 20
|||||
Db 128 ACCTTCTGGAGATCTCAGGTGTTTCATAGACCACCTGTACACCTGAAGTTCAATGTGCT 187
QY 21 LeuGlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThr 40
|||||
Db 188 CTGGGAAAGCAGCTCTGTGTGGCATTATTATGGCAGAAAGGCCTTCCCTTCCCATACT 247
QY 41 GlnPheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGln 56
|||
Db 248 CAGGTAAGTACTGACTCCACCCCTGTAGGGACCTCTTGATCTTAACCCAG 295

RESULT 4
AK047423

LOCUS

DEFINITION

AK047423

Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length

(Drosophila), full insert sequence.

AK047423

AK047423.1 GI:26338765

HTC; CAP trapper.

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

2

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Ittoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

AK047423 3447 bp mRNA linear HTC 20-SEP-2003
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:B930059M16 product:odd Oz/ten-m homolog 1
(Drosophila), full insert sequence.

AK047423

AK047423.1 GI:26338765

HTC; CAP trapper.

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

REFERENCE

2

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,

Ittoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

PUBMED
REFERENCE
AUTHORS
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3447)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. .3447
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B930059M16"
/db_xref="MGI:2412967"
/db_xref="taxon:10090"
/clone="B930059M16"
/tissue type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="10 days neonate"
862. .2472
/note="unnamed protein product; odd Oz/ten-m homolog 1
(Drosophila) (MGD|MGI:1345185, GB|NM_011855, evidence:
BLASTN, 99%, match=3037)
putative"
/codon_start=1
/protein_id="BAC33054.1"

11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
REFERENCE
AUTHORS
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3447)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
Location/Qualifiers
1. .3447
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:B930059M16"
/db_xref="MGI:2412967"
/db_xref="taxon:10090"
/clone="B930059M16"
/tissue type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev stage="10 days neonate"
862. .2472
/note="unnamed protein product; odd Oz/ten-m homolog 1
(Drosophila) (MGD|MGI:1345185, GB|NM_011855, evidence:
BLASTN, 99%, match=3037)
putative"
/codon_start=1
/protein_id="BAC33054.1"

/db_xref="GI:263338766"
/translation="MEQTDCKPYQPLSKVKHMDLAYTSSDSEDESGRKPQRQSFNSRE
TLHEYNELRRNYSQSRKRDVEKSTQIEFCETPTLCSGYHTDMHSVSRHGQLE
MGSDVDTETEGAASPDHALRMWIRGMKSEHSSCLSSRANSALSLTDTDHERKSDGENG
FKFSPVCCDMEAPADSAQDMQSPHNQFTFRPLPPPPPHACTCARPPPTVDSLQR
RSMTRSQSPSPAPAPTSTQDSVHLHNSVNLNSNIPLETRHFLFKHGSSSAIFSA
SQNYPLTSNTVYSPPPRPLRSTFRPAFTFNKPRCCNWKCTALSATAITVTLALL
AYVIVHLFTWLOQPVQIYANGISNGNPGTESMDTTYSPIGGRVSDKSEKKVFOKG
RAIDTGEVDIGAQMOTIPGFLFRFOITIHHPYILKFNISLAKDSLGIYGRNRIPP
THTQDFVKLMDGKQLVKQDSKSDDIQHSRNLILTSLOETGFIEYMDQGPWYLA
NDGKKMEQVFVLTITAIKLPCLLFEFTV"

ORIGIN

Alignment Scores: Length: 3447
Pred. No.: 2.6e-18 Matches: 37
Score: 220.50 Conservative: 20
Percent Similarity: 83.82% Mismatches: 8
Best Local Similarity: 54.41% Indels: 3
Query Match: 59.27% Gaps: 1
DB: 11
US-10-029-020-14_COPY_450_520 (1-71) x AK047423 (1-3447)
QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 2104 TTCTGGCGTTTCCAGATTACTATCCACCATCCTATATATATCTGAATTCATATTTCTTTA 2163
QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
Db 2164 GCCAAGGACTCTCTTCTGGGAATTTATGGCAGAGAAACATTCACCTACACACTCAG 2223
QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSer--- 60
Db 2224 TTTGATTTCGTGAATTAATGATGGCAACAACTGGTAAACAGGACTCTAAGAGCTCA 2283
QY 61 -----LeuGluGlyThrProArg 66
Db 2284 GATGACATTGACCACTCCCCCAAGG 2307

RESULT 5

BQ257598
LOCUS BQ257598 594 bp mRNA linear EST 06-MAY-2002
DEFINITION NISC kp04a02.q2 Baker mouse embryo e7.5 Mus musculus cDNA clone
IMAGE:5408163, mRNA sequence.
ACCESSION BQ257598
VERSION BQ257598.1 GI:20458351
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 594)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
CDNA Library Preparation: J. Baker (Stanford University)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC)
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
MGI:1844899

Seq primer: Sp6 primer.
Location/Qualifiers
1. .594
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"

FEATURES

source
row: B column: 4
Plate: LLAM12041
Seq primer: Sp6 primer.
Location/Qualifiers
1. .594
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"

CDS

/db_xref="taxon:10090"
/clone="IMAGE:5408163"
/tissue_type="embryo, late gastrula"
/dev_stage="embryo, 7.5 dpc"
/lab_host="XLI-Blue"
/clone_lib="Baker mouse embryo e7.5"
/note="Vector: pCS105; Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned into SalI/NotI sites using the following 5' adaptor: 5'-TCGACCCACGCGTCG-3'. Size-selected for average insert size 1.8-1.9 kb. Library constructed by J. Baker (Stanford University)."

ORIGIN

Alignment Scores:
Pred. No.: 5.54e-17 Length: 594
Score: 202.50 Matches: 45
Percent Similarity: 78.26% Conservative: 9
Best Local Similarity: 65.22% Mismatches: 12
Query Match: 54.44% Indels: 4
DB: 13 Gaps: 1

US-10-029-020-14_COPY_450_520 (1-71) x BQ257598 (1-594)

QY 4 ArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeuGlyLys 23
Db 9 CGATCGCAGCTCTTTATTGATCAGCCACAGTTTCTTAAAGTT-AACATCTCTTCAGAAG 67

QY 24 AlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGlnPheAsp 43
Db 68 GATGCATTGATCGGAGTGACGCCGAAGGGCTTACC GCCTTCCCATACTCAGTAGCAG 127

QY 44 PheValGluLeuAspGlyArgArgLeuThrGlnGluAlaArgSerLeu----- 61
Db 128 TTTGTGGAAC TACTGGTGTAGCAGGTTAATTGCGAGAGAGCAGCGGAACCTGTGGAG 187

QY 62 ---GluGlyThrProArgGlnSerArg 69
Db 188 TCCGAAAGAGCCGGCGGCAGCGGAGAGA 214

RESULT 6
BX494675 519 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp779B1315_r1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION DKFZp779B1315_5', mRNA sequence.

ACCESSION BX494675.1 GI:32008417

VERSION EST.
KEYWORDS Homo sapiens (human)

SOURCE
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 519)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
Unpublished (2003)
Contact: MIPS

COMMENT
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp779B1315) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..519
/organism="Homo sapiens"
/mol_type="mRNA"

/db_xref="taxon:9606"
/clone="DKFZp779B1315"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="DH10B"
/clone_lib="779 (synonym: hnccl)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB"

ORIGIN

Alignment Scores:
Pred. No.: 7.27e-17 Length: 519
Score: 201.00 Matches: 35
Percent Similarity: 77.94% Conservative: 18
Best Local Similarity: 51.47% Mismatches: 15
Query Match: 54.03% Indels: 1
DB: 13 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x BX494675 (1-519)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
Db 150 TTCTGGTT-TTCCAGATTACTATCCACCATCCCAATATATCTGAAGTTCAATATTTCTTA 208

QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
Db 209 GCCAAGGACTCTCTGCTGGGAATTTATGGCAGAGAAAACATCCACCTACATACTCAG 268

QY 42 PheAspPheValGluLeuLeuAspGlyArgArgLeuLeuThrGlnGluAlaArgSerLeu 61
Db 269 TTTGATTTTGTAAACTAATGGATGGCAACAGCTGGTCAAGCAGGACTCCAAGGGCTCT 328

QY 62 GluGlyThrProArgGlnSerArg 69
Db 329 GATGATACACAGCAGCTCCCTCGG 352

RESULT 7

CF538752

LOCUS

DEFINITION

CF538752 306 bp mRNA linear EST 12-SEP-2003
UI-M-GIO-CHF-i-06-0-UI.r1 NIH_BMAP_GIO Mus musculus cDNA clone
IMAGE:30531941 5', mRNA sequence.

ACCESSION CF538752.1 GI:34590734

VERSION EST.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 306)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

COMMENT
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: PYX-5.

Location/Qualifiers

1..306
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30531941"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GIO"

FEATURES

source

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
|||||
Db 1172 TTTTGGAGGTCAGATTCACATCAGTCAGCCTCAATTCTTAAAGTTCAACATCTCCCTG 1231

QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
|||||
Db 1232 GCGAAGGATGCCCTCTTCGGTGTCTATATAAGGAGAGGACTACCACCGTCTCATGCCCAG 1291

QY 42 PheAspPheValGluLeuLeuAspGlyArgGlyArgLeuThrGlnGluAlaArgSerLeu 61
:::|||||
Db 1292 TATGACTTCATGGAACGCCCTGGATGGAAG-----GAGAAATGGAGCGTG 1336

QY 62 GluGlyThrProArgGlnSerArg 69
:::|||||
Db 1337 GTCGAGTCGCCCCAGGGAACGCCGG 1360

RESULT 9
BM017187
LOCUS
DEFINITION BM017187 1013 bp mRNA linear EST 30-OCT-2001
603643790F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5425486 5',
mRNA sequence.
ACCESSION BM017187
VERSION BM017187.1 GI:16531541
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1013)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL
COMMENT NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM188 row: c column: 23
High quality sequence stop: 565.

FEATURES
source
1. .1013
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5425486"
/tissue_type="astrocytoma grade IV, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_98"
/note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 2.13e-14 Length: 1013
Score: 186.00 Matches: 33
Percent Similarity: 80.00% Conservative: 7
Best Local Similarity: 66.00% Mismatches: 10
Query Match: 50.00% Indels: 0
DB: 12 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x BM017187 (1-1013)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
|||||
Db 362 TTTTGGAGGTCACAAATTCACATCAGTCAGCCCCAGTTCTTAAAGTTCAACATCTCCCTC 421

QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
|||||
Db 422 GCGAAGGACGCTCTCTGTGTGTTTACATAAGAGAGGACTTCCACCATCTCATGCCCAG 481

QY 42 PheAspPheValGluLeuLeuAspGlyArg 51
:::|||||
Db 482 TATGACTTCATGGAACGCTCTGGACGGGAAG 511

RESULT 10
AL710528
LOCUS
DEFINITION AL710528 247 bp mRNA linear EST 04-SEP-2003
DKFZp686K106_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
DKFZp686K106_5', mRNA sequence.
ACCESSION AL710528
VERSION AL710528.1 GI:19693883
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 247)
JOURNAL Ansoerge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
COMMENT EST (Ansoerge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp686K106) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1. .247
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686K106"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Alignment Scores:
Pred. No.: 2.95e-13 Length: 247
Score: 171.00 Matches: 32
Percent Similarity: 78.00% Conservative: 7
Best Local Similarity: 64.00% Mismatches: 11
Query Match: 45.97% Indels: 0
DB: 9 Gaps: 0
US-10-029-020-14_COPY_450_520 (1-71) x AL710528 (1-247)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
|||||
Db 96 TTTTGAAGGTCACAAATTCACATCAGTCAGCCCCAGTTCTTAAAGTTCAACATCTCCCTC 155

QY 22 GlyLysAlaAlaLeuValGlyIleTyrGlyArgLysGlyLeuProProSerHisThrGln 41
|||||
Db 156 GCGAAGGACGCTCTCTTGGTGTGTTTACATAAGAGAGGACTTCCACCATCTCATGCCCAG 215

QY 42 PheAspPheValGluLeuLeuAspGlyArg 51


```
source
1. .392
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3893111"
/db_xref="taxon:9606"
/clone="IMAGE:285994"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Morton Fetal Cochlea"
/note="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dt. Fetal cochlea, normal. 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP XR Vector. Library constructed by N. Robertson, C. Morton. ~5' adaptor sequence: 5' GAATTCGGCAGG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN
Alignment Scores:
Pred. No.: 5.15e-09 Length: 392
Score: 142.00 Matches: 29
Percent Similarity: 85.00% Conservative: 5
Best Local Similarity: 72.50% Mismatches: 6
Query Match: 38.17% Indels: 1
DB: 14 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x N67072 (1-392)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
|||||
Db 177 TTCTGGAGATCACAGCTCTTCATTGATCAGCCACAGTTTCTTAAATTCAATATCTCTT 118
|||||

QY 22 GlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProProSerHisThrGln 41
|||||
Db 117 CAGAAGGATGCATTGATTGGAGTATATGCCGGAGAAG-TTACCGCCTTCCCACTACTCAG 59
|||||

RESULT 14
BX178536/c
LOCUS BX178536 749 bp DNA linear GSS 13-MAR-2003
DEFINITION Danio rerio genomic clone DKEY-172P5, genomic survey sequence.
ACCESSION BX178536
VERSION BX178536.1 GI:28010314
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 749)
REFERENCE Humphray,S.J., Huckle,E. and Durham,J.L.
AUTHORS Direct Submission
TITLE Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished
COMMENT This sequence was generated from the T7 end of BAC 172P5. 172P5 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details:
http://www.sanger.ac.uk/projects/D_rerio/.

FEATURES
source
1. .749
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-172P5"
/tissue_type="Testis"
/note="vector pIndigoBAC-536"

ORIGIN
Alignment Scores:
Pred. No.: 3.41e-08 Length: 749
Score: 139.00 Matches: 26
Percent Similarity: 75.00% Conservative: 4
```

```
Best Local Similarity: 65.00% Mismatches: 10
Query Match: 37.37% Indels: 0
DB: 29 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x BX178536 (1-749)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
|||||
Db 382 TTCTGGAGGTCTCTGCTTCACCTGCGCCAGCCACACTTCTCTCAAGTTCAACATCTCCCTG 323
|||||

QY 22 GlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProProSerHisThrGln 41
|||||
Db 322 GGCAAGGACGCGTGTGTTGGGGTCTACATGCGCAAGGCTGCCTCTTCACACGCTCAG 263
|||||

RESULT 15
CC122834/c
LOCUS CC122834 340 bp DNA linear GSS 16-APR-2003
DEFINITION NDL.45P17.77 Notre Dame Liverpool Aedes aegypti genomic clone
NDL.45P17, genomic survey sequence.
ACCESSION CC122834
VERSION CC122834.1 GI:29991889
KEYWORDS GSS.
SOURCE Aedes aegypti (yellow fever mosquito)
ORGANISM Aedes aegypti
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes; Stegomyia.
1 (bases 1 to 340)
REFERENCE Loftus,B., Shetty,J., Knudson,D. and Severson,D.
AUTHORS BAC end sequencing of Aedes aegypti
TITLE Unpublished (2003)
JOURNAL Other_GSSs: NDL.45P17.SP6
COMMENT Contact: Brendan Loftus
Department of Eukaryotic Genomics
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-3543
Fax: 301-838-0208
Email: enta@tigr.org
Library was provided by David Severson
Seq primer: T7
Class: BAC ends.

FEATURES
source
1. .340
/organism="Aedes aegypti"
/mol_type="genomic DNA"
/strain="liverpool"
/db_xref="taxon:7159"
/clone="NDL.45P17"
/clone_lib="Notre Dame Liverpool"
/note="Vector: pECBAC1; Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN
Alignment Scores:
Pred. No.: 1.05e-05 Length: 340
Score: 117.00 Matches: 21
Percent Similarity: 61.54% Conservative: 11
Best Local Similarity: 40.38% Mismatches: 20
Query Match: 31.45% Indels: 0
DB: 28 Gaps: 0

US-10-029-020-14_COPY_450_520 (1-71) x CC122834 (1-340)

QY 2 PheTrpArgSerGlnValPheIleAspHisProValHisLeuLysPheAsnValSerLeu 21
|||||
Db 250 TTTTGAACGCAGCAATTAGTAACAACATCTGTCATTCAATTTAATTTCACTTTA 191
|||||

QY 22 GlyLysAlaAlaLeuValGlyIleTyrglyArgLysGlyLeuProProSerHisThrGln 41
|||||
Db 190 CCGTGGGGTGCCAACTTTGCTGTGTACGGTCACGAAATGTAGGACCTAGCATAACGCAA 131
|||||
```

Qy 42 PheAspPheValGluLeuLeuAspGlyArgArgLeu 53
Db 130 TATGATTTTGTAGAAATTATATAAAGGTGGAAGGATA 95

Search completed: August 14, 2004, 18:02:47
Job time : 842.674 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 14, 2004, 01:13:26 ; Search time 4691.12 Seconds
(without alignments)
2546.273 Million cell updates/sec

Title: US-10-029-020-14_COPY_1_400
Perfect score: 2143
Sequence: 1 MDVKERKPYRSLTRRRDAER.....EITEDTASSWPVPTDVSLYP 400

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US10029020/runat_06082004_112216_29287/app_query.fasta_1.3519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10029020@cgn_1_113135@runat_06082004_112216_29287 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	936	43.7	620	14	CB532245	CB532245 756697 MA
2	936	43.7	620	14	CB532612	CB532612 757081 MA
3	884	41.3	566	13	BU072782	BU072782 im47g02.Y
4	781	36.4	625	13	BU200362	BU200362 604158648
5	737	34.4	878	13	BU474934	BU474934 603364850
6	722.5	33.7	3447	11	AK047423	AK047423 Mus muscu
7	699	32.6	884	13	BQ735622	BQ735622 AGENCOURT
8	692	32.3	789	13	BU474512	BU474512 603760992
9	687.5	32.1	887	13	BU119163	BU119163 603143046
10	685	32.0	402	14	CB544750	CB544750 AMGNNUC:T
11	659	30.8	4556	11	AK034286	AK034286 Mus muscu
C 12	596.5	27.8	490	9	AL922332	AL922332 AL922332
13	568	26.5	870	10	BG036207	BG036207 602326960
14	531.5	24.8	511	29	AY405291	AY405291 Homo sapi
15	531.5	24.8	511	29	AY405292	AY405292 Pan trogl
16	529.5	24.7	511	29	AY405293	AY405293 Mus muscu
17	524.5	24.5	534	10	BF944080	BF944080 QV3-NN020
18	509	23.8	2716	11	AK031198	AK031198 Mus muscu
C 19	502.5	23.4	880	13	BX760712	BX760712 BX760712
C 20	482.5	22.5	568	10	BF953105	BF953105 QV3-NN020
21	470	21.9	2627	11	AK050784	AK050784 Mus muscu
22	456.5	21.3	643	10	BB657984	BB657984 BB657984
23	448	20.9	717	10	BB654584	BB654584 BB654584
C 24	438.5	20.5	593	29	CE756467	CE756467 tigr-gss-
C 25	433	20.2	284	10	BF388169	BF388169 UI-R-CA1-
26	417	19.5	259	28	CC178145	CC178145 XC549 Bay
27	409	19.1	825	13	BU253342	BU253342 603747714
28	388.5	18.1	843	14	CF225019	CF225019 AGENCOURT
29	381	17.8	864	13	BU172491	BU172491 AGENCOURT
30	374	17.5	397	13	BY010856	BY010856 BY010856
31	374	17.5	704	13	BY733523	BY733523 BY733523
32	367	17.1	643	13	BY723994	BY723994 BY723994
33	343.5	16.0	608	9	AL955817	AL955817 AL955817
34	330.5	15.4	707	13	BU229930	BU229930 603947544
35	329.5	15.4	801	13	BU451304	BU451304 603216940
36	295	13.8	461	13	BU473055	BU473055 603761665
37	282.5	13.2	404	29	CG514853	CG514853 OST68836
38	282.5	13.2	751	13	BU461850	BU461850 603773816
C 39	273.5	12.8	1060	29	CNS04V97	AL308788 Tetraodon
40	272.5	12.7	745	29	CE434866	CE434866 tigr-gss-
C 41	257.5	12.0	386	9	AA702693	AA702693 zi90g04.s
42	256.5	12.0	674	28	BZ111628	BZ111628 CH230-230
C 43	248	11.6	474	9	AI753786	AI753786 crl6e09.x
44	247.5	11.5	600	13	BU921705	BU921705 6093-12 M
45	247.5	11.5	644	10	BB657539	BB657539 BB657539

ALIGNMENTS

RESULT 1 CB532245/c
LOCUS CB532245 620 bp mRNA linear EST 16-MAY-2003
DEFINITION 756697 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
ACCESSION CB532245
VERSION CB532245.1 GI:29395750
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 620)

Db 92 AACACGAGACCGATCACCCGGGGCGCTGCAGAACCATTTCTCGACTCCGGACACCGCGC 33

QY 181 ProProLeuSerHisAlaHisThrPro 189
|||||

Db 32 CCTCCGCTCTCGCATGCCACACACCCC 6

RESULT 2
CB532612 620 bp mRNA linear EST 16-MAY-2003
LOCUS 757081 MARC 6BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION CB532612
ACCESSION CB532612
VERSION CB532612.1 GI:29396493
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 620)
AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: LAM8014 row: K column: 19
Seq primer: GTAATACGACTCACTATAGGG.
Location/Qualifiers
1..620
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

FEATURES
source
14

ORIGIN
Alignment Scores: 2.05e-58 Length: 620
Pred. No.: 936.00 Matches: 174
Score: 95.77% Conservative: 7
Percent Similarity: 92.06% Mismatches: 8
Best Local Similarity: 43.68% Indels: 0
Query Match: 14 Gaps: 0
DB:

US-10-029-020-14_COPY_1_400 (1-400) x CB532612 (1-620)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
|||||

Db 49 ATGGACGTGAAGGAGAGAGGAGCCATACCGCTCGCTGACCCGGCGCGGACCGCGCGC 108
|||||

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
|||||

Db 109 CGCTACACCAAGCTCGTCGGCCGACAGCGAAGAGGGCAAGGCCCGCGAGTCTCTACAGC 168
|||||

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||||

Db 169 TCCAGCGAGACCTCAAGGCCTACGACCGAGCGCGCCTCACCTACGCGCGCGTC 228
|||||

QY 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
|||||

Db 229 AAGGACCTGGTACACAGGAGGCGGAGGAGTTCTGCGCGCGAGCCACCACTTCAGCCTG 288
|||||

AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G.,
Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: LAM8014 row: K column: 19
Seq primer: TAGAAGGCACAGTCGAGG.
Location/Qualifiers
1..620
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including liver, lung, hypothalamus, pituitary, and
placenta/endometrium."

FEATURES
source
14

ORIGIN
Alignment Scores: 2.05e-58 Length: 620
Pred. No.: 936.00 Matches: 174
Score: 95.77% Conservative: 7
Percent Similarity: 92.06% Mismatches: 8
Best Local Similarity: 43.68% Indels: 0
Query Match: 14 Gaps: 0
DB:

US-10-029-020-14_COPY_1_400 (1-400) x CB532245 (1-620)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
|||||

Db 572 ATGGACGTGAAGGAGAGGAGCCATACCGCTCGCTGACCCGGCGCGGACCGCGCGC 513
|||||

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
|||||

Db 512 CGCTACACCAAGCTCGTCGGCCGACAGGAGGCGGAGGCGCGGAGTCTGCGCGCGAGCTCTACAGC 453
|||||

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
|||||

Db 452 TCCAGCGAGACCTCAAGGCCTACGACCGAGCGCGCTCACTTACGCGCGCGCTC 393
|||||

QY 61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
|||||

Db 392 AAGGACCTGGTACACAGGAGGCGGAGGAGTCTGCGCGCGAGCTCTGAGGCGGAGCGGAC 333
|||||

QY 81 ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle 100
|||||

Db 332 CGCGAGCTGGGGCTGGGGAGGTGACACCCCGCACGGACGCTGTACCGCACGGACATC 273
|||||

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
|||||

Db 272 GGCCTGCCCCACTCGGCTACTCCCTGGGCGCCAGCTCTGAGGCGGAGCTGGAGGCGGAC 213
|||||

QY 121 ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg 140
|||||

Db 212 GCCGCACTGTCCCCCGAGCACCCCGTGGCGGTGTGGGCGCGCAGCACCGGTTCAGGACGC 153
|||||

QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
|||||

Db 152 AGCTCCTGCCTGTCAGCGCGGCCAACCTCCAACTCACGCTCACTGACACGAGCACGAG 93
|||||

QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
|||||

QY 81 ArgGluLeuGlyLeuGluValThrProHisGlyThrLeuTyArgThrAspIle 100
Db 289 CGCGAGCTGGGCTGGGGAGGTGACACCCCGCACGGGACGCTGTACCGCACGGACATC 348
QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 349 GGCCTGCCCACTGGCGCTACTCCCTGGGCGCCAGCTCTGAGGCCGAGCTGGAGCGGCAG 408
QY 121 ThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArgSerGlyArg 140
Db 409 GCCGCACTGTCCCCGAGCACCCCGTGTGGGCGCGCAGCACCGCGGTGAGGACGC 468
QY 141 SerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGlu 160
Db 469 AGCTCTGCCTGTCCAGCCGGGCCAACTCCAACTCACGCTCACTGACACGGAGCACGAG 528
QY 161 AsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProPro 180
Db 529 AACACGGAGACCGATCACCCCGCGCGCTGCAGAACCATTTCTCGACTCCGGACACCGCCG 588
QY 181 ProProLeuSerHisAlaHisThrPro 189
Db 589 CCTCGCTCTCGCATGCCACACACCCCC 615

RESULT 3
BU072782
LOCUS
DEFINITION
im47g02.y1 HR85 islet Homo sapiens cDNA clone IMAGE:6038211 5',
similar to TR:O14667 O14667 PRO-NEUREGULIN-1, GAMMA ISOFORM
PRECUSOR ;, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 566)

REFERENCE
AUTHORS
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)

TITLE
JOURNAL
COMMENT
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)

Seq primer: -40RP from Gibco
High quality sequence stop: 339.
Location/Qualifiers
1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6038211"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_lib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:
NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'

XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."

ORIGIN

Alignment Scores:
Pred. No.: 1.12e-54 Length: 566
Score: 884.00 Matches: 162
Percent Similarity: 95.86% Conservative: 0
Best Local Similarity: 95.86% Mismatches: 7
Query Match: 41.25% Indels: 0
DB: 13 Gaps: 0
US-10-029-020-14_COPY_1_400 (1-400) x BU072782 (1-566)
QY 165 AspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArgThrProProProLeuSer 184
Db 57 GATCATCCGGCGGCGCTGCAGAACACACGCGCGGCTCCGGACGCGCGCGCTCTCG 116
QY 185 HisAlaHisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsn 204
Db 117 CACGCCACACACCCCCAACCCAGACCCACGCGGCTCCATTAACTCCCTGAACCGGGCAAC 176
QY 205 PheThrProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluPro 224
Db 177 TTCACGCGGAGGAGCAACCCAGCCCGGCCCCACCGGACCACTCGCTCTCCGGAGAGCCC 236
QY 225 ProAlaGlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsn 244
Db 237 CTGCGCGGCGGCCAGGAGCCTGCCACGCCAGGAGAACTGGCTGCTCAACAGCAAC 296
QY 245 IleProLeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsn 264
Db 297 ATCCCCCTGGAGACCAGAAACCTAGGCAAGCAGCATTCCTAGGACATTGCAGGACAAC 356
QY 265 LeuIleGluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHis 284
Db 357 CTCATTGAGATGGACATTCTCGGCGCCTCCCGCCATGATGGGCTTACAGTGACGGGCAC 416
QY 285 PheLeuPheLysProGlyGlyThrSerProLeuPheCysThrThrSerProGlyTyrPro 304
Db 417 TTCTCTTTCAAGCCTGGAGGCACCTCCCCGCTCTTCTGCACCACATCACCAGGTACCCA 476
QY 305 LeuThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAla 324
Db 477 CTGACGTCCAGCACAGTGACTCTCCCTCCCGCCGACCCCTGGCCCGAGAACTTTTCGC 536
QY 325 ArgProAlaPheAsnLeuLysLysPro 333
Db 537 CGGGCGGNCCTTAACCTCAAGAGCCT 563

RESULT 4

BU200362

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 625)

Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,

Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

A Comprehensive Collection of Chicken cDNAs

Curr. Biol. 12 (22), 1965-1969 (2002)

22335534

BU200362 625 bp mRNA linear EST 25-NOV-2002
604158648F1 CSEQCHN03 Gallus gallus cDNA clone ChEST1010p11 5',
mRNA sequence.

BU200362
BU200362.1 GI:25363728
EST.
Gallus gallus (chicken)

Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

1 (bases 1 to 625)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)

PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .625
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChEST1010p11"
/tissue_type="whole embryo"
/dev_stage="20-21"
/lab_host="DH10B"
/clone_lib="CSEQCHN03"
/note="Organ: whole embryo; Vector: pBluescript II KS(+);
Site_1: EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN
Alignment Scores:
Pred. No.: 4.25e-47 Length: 625
Score: 781.00 Matches: 151
Percent Similarity: 91.62% Conservative: 13
Best Local Similarity: 84.36% Mismatches: 13
Query Match: 36.44% Indels: 2
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x BU200362 (1-625)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db 93 ATGGATGTAAGAAAGAAACCGTATCGATCTCTGACTCGGCGCGGACACGGAGCGC 152

QY 21 ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db 153 CGCTACACAGCTCTTCAGCCGAGAGTGAGGACAGCAAGGCTCCTCAGAACTCTATAGC 212

QY 41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db 213 TCCAGTGAGACCTGAAGGCTTATGATCAAGACTCCAGGTTGACCTACAGCAATCGGGTC 272

QY 61 LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db 273 AAAGACATGGTGACACAGGAGGCTGATGATCTGCGGAGCAGGAGCCCACTTCTCTTG 332

QY 81 ArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArgThrAspIle 100
Db 333 CGGGAGCTGGGTCTTGAAGATGTGACTCCACCCATGGGACTTTGTACCGGACTGATATT 392

QY 101 GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp 120
Db 393 GGGCTGCCTCACTGTGGCTACTCCATCAGCGCTGGCTCAGATGCCGACACGGAGCGGAT 452

QY 121 ThrValLeuSerProGluHis-ProValArgLeuTrpGlyArgSerThrArgSerGlyAr 140
Db 453 GTGGTCATGTACCTGAGCATCCCTGTGAGGCTCTGGGACGCAACACCAAAATCCGACG 512

QY 140 gSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1 160
Db 513 CAGCTCCTGCTTGTGAGCGCGGCAACTCCAACTCACCTTACCGACACGGAGCACA 572

QY 160 uAsn-ThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg 177
Db 573 GAATTACCGAAACTGATCATCTCTCCAAATCTTCAAATATCATTCAGACTCCGA 625

RESULT 5
BU474934 878 bp mRNA linear EST 30-NOV-2002
LOCUS 603364850F1 CSEQRBN21 Gallus gallus cDNA clone ChEST262c14 5', mRNA
DEFINITION sequence.
BU474934
BU474934.1 GI:25968511
EST.
Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 878)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers
1. .878
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="ChEST262c14"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBN21"
/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

FEATURES
source
1. .878

ORIGIN
Alignment Scores:
Pred. No.: 1.13e-43 Length: 878
Score: 737.00 Matches: 141
Percent Similarity: 92.73% Conservative: 12
Best Local Similarity: 85.45% Mismatches: 11
Query Match: 34.39% Indels: 1
DB: 13 Gaps: 0

US-10-029-020-14_COPY_1_400 (1-400) x BU474934 (1-878)

NDGKMEQVFVLTTAIGKLPCLLFEFTV"

ORIGIN

Alignment Scores: 9.2e-42 Length: 3447

Pred. No.: 722.50 Matches: 169

Score: 53.01% Conservativity: 51

Best Local Similarity: 40.72% Mismatches: 100

Query Match: 33.71% Indels: 95

DB: 11 Gaps: 15

US-10-029-020-14_COPY_1_400 (1-400) x AK047423 (1-3447)

QY 1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19

DB 862 ATGGAGCAACACAGACTGCAACCTTATCAGCCTCTGTCCAAAGTCAAGCATGAATGGAT 921

QY 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyr 39

DB 922 CTAGCTTATACCAAGTTCTCTGATGAGAGTGAAGATGGGAGAAACCAAGACAGTCATTC 981

QY 40 SerSerSerGluThrLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer--- 58

DB 982 AACTCCAGGAACTCTGCATGAGTATACCAAGAGTCTGAGGAGGAATTACAAATAGCCAG 1041

QY 59 -----ArgValLysAspIleValProGlnGluAlaGlu---GluPheCysArgThrGly 75

DB 1042 ACTAGAAAGAGGAAGATGGAGAAATCTACTCAAGAGATAGATTCTGTGAA----- 1095

QY 76 AlaAsnPheThrLeuArgGluLeuGlyLeuGluValThrProProHisGlyThrLeu 95

DB 1096 -----ACACCTCCT-----ACTTTG 1110

QY 96 -----TyrArgThrAspIle---GlyLeuProGlnCysGlyTyrSerMetGlyAla 111

DB 1111 TGCTCCGGCTACACACAGACATGCACAGTGTCTTCTCGACATGGCTACCAGCTGGAGATG 1170

QY 112 GlySerAspAlaAspMetGluAlaAspThrValLeuSerProGluHisProValArgLeu 131

DB 1171 GGATCTGATGATAGATACAGACAGAGAGGAGTGCATCACCTGACCATGCATCAGAATG 1230

QY 132 TrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsn 151

DB 1231 TGGATAAGAGGCATGAATCAGAACACAGTCTCTGTCTCCAGTAGGGCCAACTCTGCA 1290

QY 152 LeuThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeuGln 171

DB 1291 CTGTCTTTCACCGACACTGATCATGAACGG--AAGTCTGATGGGGAATAATGTTTAA 1347

QY 171 ----- 171

DB 1348 TTCTCTCTGTTTGTGTGACATGGAGGCTCCAGCTGATTTCGGCTCAAGACATGCAAAGC 1407

QY 172 -----AsnHisAlaArgLeuArg-----ThrProProProLeuSerHisAla 186

DB 1408 AGCCCAACACAAACAGTTTACCTTCAGACCCCTCCACCCACCCGACCTCTCCACATGCC 1467

QY 187 HisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206

DB 1468 TGCACCTGTGCCAGGAAGCCACCTCCTACAGTGGACTCTCTACAAAGAACATCAATGACT 1527

QY 207 ProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProProAla 226

DB 1528 ACCCGCAGCCAGCCCGAGCCAGCT-----GCTCTGCTCCTCCA 1566

QY 227 GlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIlePro 246

DB 1567 ACCAGCACACAGGATTTCAGTTTCATCTGCATACAGCTGGGTCTTGAACAGTAACATACCA 1626

QY 247 LeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIle 266

DB 1627 CTGGAGACCAGG----- 1638

QY 267 GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeu 286

DB 1639 -----CATTTCCTG 1647

QY 287 PheLysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305

DB 1648 TTCAAACATGGATCTGTTCTTCTGCTATCTTCTAGTGCAGCCAGTCAAGCTCTG 1707

QY 306 ThrSerSerThrValTyrSerProProProArgProLeuProArgSerThrPheAlaArg 325

DB 1708 ACATCTAATACTGTCTACTCACCACCAACCAGCGCTGCCTCGAAGCACCTTTCCCGA 1767

QY 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345

DB 1768 CCTGCCTTCACTTTTAAACAAACCATACAGATGCTGCAATTGGAAGTGCACAGCCTGAGC 1827

QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeu 365

DB 1828 GCCACTGCAATCACAGTACTTTGGCCTTGTACTAGCCTAT---GTAATTGTACACTTG 1884

QY 366 PheGlyLeuAsnTrpHisLeuGlnProMetGluGlyGlnMetTyr 380

DB 1885 TTTGGCCTGACGTGGCAGTTGCAACCAGTT---GGACAGATCTAT 1926

RESULT 7

BQ735622 884 bp mRNA linear EST 16-JUL-2002

LOCUS AGENCOURT_8097117 NICHD_XGC_Emb4 Xenopus laevis cDNA clone

DEFINITION IMAGE:5542635 5', mRNA sequence.

ACCESSION BQ735622

VERSION BQ735622.1 GI:21874519

KEYWORDS EST.

SOURCE Xenopus laevis (African clawed frog)

ORGANISM Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopus; Xenopodinae; Xenopus.

REFERENCE 1 (bases 1 to 884)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Plate: LLAM12242 row: e column: 04
High quality sequence stop: 373.

FEATURES

Location/Qualifiers

1..884

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:5542635"

/dev_stage="embryo, stage 31-32"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NICHD_XGC_Emb4"

/note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2.1 kb. Constructed by Life Technologies. Note: This is a Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores: 6.75e-41 Length: 884

Pred. No.: 699.00 Matches: 160

Score: 60.76% Conservativity: 15

Best Local Similarity: 55.56%		Mismatches: 44
Query Match: 32.62%		Indels: 69
DB: 13		Gaps: 8
US-10-029-020-14_COPY_1_400 (1-400) x BQ735622 (1-884)		
QY	113 SerAspAlaAspMetGluAlaAspThrValLeuSerProGluHisProValArgLeuTrp	132
Db	26 AACGATTCTGAGACTCTGAAAGAGACTTTGTGTC-----	55
QY	133 GlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeu	152
Db	56 -----CCTAATTCCTTCTTAGTTAAAGTGATTCTGCACAACTG	94
QY	153 ThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeuGlnAsn	172
Db	95 GCTGTC-----CCTGCTAACGATCAGCCGCTGTTCTGCAGAAAT	133
QY	173 HisAlaArgLeuArgThrProProProProLeuSerHisAlaHisThrProAsnGlnHis	192
Db	134 CATTGTGAGACTACGAACGCCCCACCGCCCTTAAGCCACCCGAAACCC-----CAC	187
QY	193 HisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThrProArgSerAsnProSer	212
Db	188 CATGCGGCATCCATCAATTCTTTGAACCGGGCAATTACACCCACGTAGCAACCTAGC	247
QY	213 ProAlaProThrAspHisSerLeuSerGlyGluProProAlaGlyGlyAlaGlnGluPro	232
Db	248 CCGCGCCCGACGGACCACTCTCTGTGGGTGAGCCCCCG---GGCGGTGCCAGGAGTCC	304
QY	233 AlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIleProLeuGluThrArgAsnLeu	252
Db	305 GTCCAT-----GACAATTGGCTGTGAACAGCAACATTCCCTGGAGACCAGG-----	352
QY	253 GlyLysGlnPropheLeuGlyThrLeuGlnAspAsnLeuIleGluMetAspIleLeuGly	272
Db	352 -----	352
QY	273 AlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeupheLysProGlyGlyThr	292
Db	353 -----CATTTTCTGTTTAAGCCTGGATGACG	379
QY	293 SerProLeupheCysThrThrSerProGlyTyrProLeuThrSerSerThrValTyrSer	312
Db	380 TCCCTCTGTTCTGCACGACGTCTCCTGGAAACCCCTGACATCCAGCACTGTGTACTCC	439
QY	313 ProProProArgProLeuProArgSerThrPheAlaArg-ProAlaPheAsnLeuLysLy	332
Db	440 CCTCCTCCGAGGACGCTTCCTCGCAGCACTTCTCGCGGGTCGGGCTTTAACCTTAAGAA	499
QY	332 sProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSerAlaIleValIleSerAlaTh	352
Db	500 GCCGTACAAATACTGCAACTGGAAATGTGCCGCCCTGAGCGCCATCATCATCTCCGTAC	559
QY	352 rLeuValIleLeuLeuAla-TyrPheValAlaMetHisLeu--PheGlyLeuAsnTrp--	370
Db	560 TCTGAAGATCCTCTGGGATTATTTTCATCGCCATGCACCTTGTGTTGGGCTAAACTGGGC	619
QY	371 -HisLeuGlnProMetGluGlyGlnMetTyrGlu-IleThrGlu---AspThrAlaSerS	389
Db	620 CCCTGCTGCCGAACGAAGGGCCAGAGTATGACACTCCCGGAAGGACAGAGGGCGGCAA	679
QY	389 erTrpProValPro 393	
Db	680 TTTGGCCACGTCCC 693	
RESULT 8		
BU474512		
LOCUS		
DEFINITION		
603760992F1 CSEQRBN21 Gallus gallus cdna clone ChEST679g24 5', mRNA		
sequence.		
ACCESSION		
BU474512		
VERSION		
BU474512.1 GI:25968089		

KEYWORDS	EST.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
REFERENCE	Phasianinae; Gallus.
AUTHORS	1 (bases 1 to 789)
	Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
	Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE	A Comprehensive Collection of Chicken cDNAs
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE	22335534
PUBMED	12445392
COMMENT	Contact: Simon Hubbard
	Department of Biomolecular Sciences
	University of Manchester Institute of Science and Technology
	(UMIST)
	PO Box 88, Manchester, M60 1QD, UK
	Tel: 01612008930
	Fax: 01612360409
	Email: Simon.Hubbard@umist.ac.uk.
FEATURES	Location/Qualifiers
source	1..789
	/organism="Gallus gallus"
	/mol_type="mRNA"
	/strain="Layer"
	/db_xref="taxon:9031"
	/clone="ChEST679g24"
	/sex="Female"
	/dev_stage="adult"
	/lab_host="DH10B"
	/clone_lib="CSEQRBN21"
	/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
	EcoRI; Site_2: NotI; This normalized library was
	constructed from 1 million independent clones. cDNA
	synthesis was initiated using an oligo(dT) primer, using
	methyalted C in the first strand synthesis reaction.
	Following this first strand reaction, double-stranded cDNA
	was bluntend, ligated to NotI adapters, digested with
	EcoRI, size-selected, and cloned into the NotI and EcoRI
	compatible sites of a custom modified MCS of the
	pBluescript (KS+) vector. The library was normalized in 2
	rounds using conditions adapted from Soares et al., PNAS
	(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
	(1996): 791, except that a significantly longer
	reannealing hybridization was used."

ORIGIN

Alignment Scores:	
Pred. No.:	1.86e-40
Score:	692.00
Length:	789
Percent Similarity:	91.52%
Matches:	139
Best Local Similarity:	84.24%
Conservative:	12
Mismatches:	13
Query Match:	32.29%
Indels:	3
DB:	13
Gaps:	0
US-10-029-020-14_COPY_1_400 (1-400) x BU474512 (1-789)	
QY	1 MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAspAlaGluArg 20
Db	256 ATGGATGTAAGAAAGAAAGAAACCGTATCGATCTCTGACTCGGCGCCGACACGGAGCGC 315
QY	21 ArgTyrThrSerSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer 40
Db	316 CGCTACACCACTCTTCAGCCCGAGAGTGAGGACAGCAAGGCTCCTCAGAAGTCTATAGC 375
QY	41 SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal 60
Db	376 TCCAGTGAGACCCCTGAAGGCTTATGATCAAGACTCCAGGTTCACCTACAGCAATCGGTC 435
QY	61 LysAspIleValProGlnGluAlaGluPheCysArgThrGlyAlaAsnPheThrLeu 80
Db	436 AAAGACATGGTGCACCCAGGAGCTGATGAATTCTGCCGAGCAGGAGGCCAATCTCTTTG 495

Alignment Scores:		4.68e-40	Length:	887
Pred. No.:		687.50	Matches:	146
Score:		87.43%	Conservative:	14
Percent Similarity:		79.78%	Mismatches:	21
Best Local Similarity:		32.08%	Indels:	5
Query Match:		13	Gaps:	0
DB:				
US-10-029-020-14_COPY_1_400 (1-400) x BUI19163 (1-887)				
QY	1	MetAspValLysGluArgLysProTyrArgSerLeuThrArgArgAlaGluArg	20	
Db	343	ATGGATGTAAAGAAAGAAACCGTATCGATCTCTGACTCGGCGCGGACACGGAGCGC	402	
QY	21	ArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSer	40	
Db	403	CGCTACACCAGCTCTTCAGCCGAGAGTGAGGACAGCAAGGCTCCTCAGAAATCCTATAGC	462	
QY	41	SerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgVal	60	
Db	463	TCCAGTGAGACCCCTGAAGGCTTATGATCAAGACTCCAGGTTGACCTACAGCAATCGGGTC	522	
QY	61	LysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhetheLeu	80	
Db	523	AAAGACATGGTGCCACCAGGAGGCTGATGATTTCTGCCGAGCAGGAGCCAACTTCTCTTTG	582	
QY	81	ArgGluLeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100	
Db	583	CGGAGCTGGGTCTTGAAGATGTGACTCCCACTGGGACTTTGTACCGGACTGATATT	642	
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120	
Db	643	GGGCTGCCTCACTGTGGCTACTCCATCAGCGCTGGCTCAGATGCCACACGGAGCGGAT	702	
QY	121	ThrValLeuSer-ProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyAr	140	
Db	703	GTGGTCATGTCACCTTGAGCATCTCTGTGAGGCTCTGGGACCGCAACCAATCCGGACG	762	
QY	140	gSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1	160	
Db	763	CAGCTCCTGCTGTGAGCGCGGCCAACTCAACTC-ACCCTCACCAGACACGGAAGCCCG	821	
QY	160	uAsnThrGluThrAspHisProGlyGlyLeuGlnAsnHisAlaArgLeuArg-ThrProp	180	
Db	822	AGAATACGAAAT-GATCATCTCTCCATATCTTCA-AATCATTCAAGACTCCGAATCCACAC	879	
QY	180	roPro 181		
Db	880	CTCCA 884		
RESULT 10				
LOCUS	CB544750	402 bp	mRNA	linear
DEFINITION	AMGNNUC:TRYP1-00009-A5-A tryP1 (10582) Rattus norvegicus cDNA clone			
ACCESSION	CB544750			
VERSION	CB544750.1	GI:29428647		
KEYWORDS	EST.			
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
REFERENCE	1 (bases 1 to 402)			
AUTHORS	Amgen EST Program.			
TITLE	Amgen Rat EST Program			
JOURNAL	Unpublished (2003)			
COMMENT	Contact: Dan Fitzpatrick Amgen, Inc One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA Tel: 805 447-4881 Plate: 00009 row: a column: 5.			
FEATURES				
Location/Qualifiers				
QY	180	roPro 181		
Db	880	CTCCA 884		
RESULT 9				
LOCUS	BUI19163	887 bp	mRNA	linear
DEFINITION	603143046F1 CSEQCHL16 Gallus gallus cDNA clone ChEST137111 5', mRNA			
ACCESSION	BUI19163			
VERSION	BUI19163.1	GI:25328007		
KEYWORDS	EST.			
SOURCE	Gallus gallus (chicken)			
ORGANISM	Gallus gallus			
REFERENCE	1 (bases 1 to 887)			
AUTHORS	Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.			
TITLE	A Comprehensive Collection of Chicken cDNAs			
JOURNAL	Curr. Biol. 12 (22), 1965-1969 (2002)			
MEDLINE	22335534			
PUBMED	12445392			
COMMENT	Contact: Simon Hubbard Department of Biomolecular Sciences University of Manchester Institute of Science and Technology (UMIST) PO Box 88, Manchester, M60 1QD, UK Tel: 01612008930 Fax: 01612360409 Email: Simon.Hubbard@umist.ac.uk.			
FEATURES				
source				
QY	81	ArgGluLeuGlyLeuGluValThrProProHisGlyThrLeuTyrArgThrAspIle	100	
Db	496	CGGAGCTGGTCTTGAAGATGTGACTCCACCCATGGGACTTTGTACCGGACTGATATT	555	
QY	101	GlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAsp	120	
Db	556	GGGCTGCCTCACTGTGGCTACTCCATCAGCGCTGGCTCAGATGCCGACACGGAGCGGAT	615	
QY	121	ThrValLeuSerProGluHisProValArgLeuTyrGlyArgSerThrArgSerGlyArg	140	
Db	616	GTGGTCATGTACCTGAGCATCTGTGAGGCTCTGGGACGCAACACCAATCCGACGC	675	
QY	141	SerSer-CysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisG1	160	
Db	676	AGCTCCCTGCTTGTGAGCGCGGC-AACTCCAACCTCACCTC-ACCGACACGGAGCAGCA	733	
QY	160	uAsnThrGluThr 164		
Db	734	GAATACCGAAACT 746		
ORIGIN				
[5'aattcttttttcggtatccgggggtgcacgc]				

ORIGIN

Alignment Scores: 5.91e-37 Length: 4556
Pred. No.: 659.00 Matches: 163
Score: 49.65% Conservative: 52
Percent Similarity: 37.64% Mismatches: 109
Best Local Similarity: 30.75% Indels: 109
Query Match: 14 Gaps: 14
DB:

US-10-029-020-14_COPY_1_400 (1-400) x AK034286 (1-4556)

QY 1 MetAspVallysGluArgLysProTyrArgSerLeuThrArg--ArgArgAspAlaGlu 19
Db 987 ATGGAGCAACAGACTGCAAACTTATCAGCCTCTGTCCAAAGTCAAGCATGAATGGAT 1046
QY 20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyr 39
Db 1047 CTAGCTTATACCAGTTCTTCTGATGAGAGTGAAGATGGGAGAAACCAAGACAGTCAATC 1106
QY 40 SerSerSerGluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySer-- 58
Db 1107 AACTCCAGGAAACTCTGCATGATGAGATGAAGATGGGAGAAACCAAGACAGTCAATC 1166
QY 59 -----ArgVallysAspIleValProGlnGluAlaGlu---GluPheCysArgThrGly 75
Db 1167 AGTAGAAGAGGAAAGATGTGGAGAAATCTACTCAAGAGATAGAAATCTGTGAA----- 1220
QY 76 AlaAsnPheThrLeuArgGluLeuGluGluValThrProHisGlyThrLeu 95
Db 1221 -----ACACCTCCT-----ACTTTG 1235
QY 96 -----TyrArgThrAspIle---GlyLeuProGlnCysGlyTyrSerMetGlyAla 111
Db 1236 TGCTCCGGCTACACACAGACATGCACAGTGTCTCGACATGGCTACAGCTGGAGATG 1295
QY 112 GlySerAspAlaAspMetGluAlaAspThrValLeuSerProGluHisProValArgLeu 131
Db 1296 GGATCTGATGTAGATACAGAGACAGAGGAGGCTGCATCACCTGACCATGCATCAGAATG 1355
QY 132 TrpGlyArgSerThrArgSerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsn 151
Db 1356 TGGATAAGAGGAGCATGAATCAGAAATCAGAAACACAGTTCCTGTCTGTCCAGTAGGGCCCACTGTGCA 1415
QY 152 LeuThrLeuThrAspThrGluHisGluAsnThrGluThrAspHisProGlyGlyLeuGln 171
Db 1416 CTGTCTTGACCGACACTGATCATGAACGG---AAGTCTGATGGGAAATGGTTTAAA 1472
QY 171 ----- 171
Db 1473 TTCTCTCCTGTTTGTGTGACATGAGGCTCCAGCTGATCGGCTCAAGACATGCAAGC 1532
QY 172 -----AsnHisAlaArgLeuArg-----ThrProProProLeuSerHisAla 186
Db 1533 AGCCACACAAACCAGTTACCTTCAGACCCCTCCACCCAGCCACCTCCTCCACATGCC 1592
QY 187 HisThrProAsnGlnHisHisAlaAlaSerIleAsnSerLeuAsnArgGlyAsnPheThr 206
Db 1593 TGCACCTGTGCCAGGAAGCCACCTCTACAGTGGACTCTCTACAAAGAGATCAATGACT 1652
QY 207 ProArgSerAsnProSerProAlaProThrAspHisSerLeuSerGlyGluProAla 226
Db 1653 ACCCGCAGCCAGCCAGCCAGCT-----GCTCCTGCTCCTCCA 1691
QY 227 GlyGlyAlaGlnGluProAlaHisAlaGlnGluAsnTrpLeuLeuAsnSerAsnIlePro 246
Db 1692 ACCAGCACACAGGATTCAGTTCTCTGCTAAGAGCTGGGTCTTGAACAGTAACATACCA 1751
QY 247 LeuGluThrArgAsnLeuGlyLysGlnProPheLeuGlyThrLeuGlnAspAsnLeuIle 266
Db 1752 CTGGAGACCCAGG----- 1763
QY 267 GluMetAspIleLeuGlyAlaSerArgHisAspGlyAlaTyrSerAspGlyHisPheLeu 286

Db 1764

QY 287 PheLysProGly---GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeu 305
Db 1773 TTCAAACATGGATCGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1832
QY 306 ThrSerSerThrValTyrSerProProProProProProProProProProProProPro 325
Db 1833 ACATCTAATACTGTCTACTCACCACCACCAGCGCGCTGCTCGAAGCACCTTTTCCCGA 1892
QY 326 ProAlaPheAsnLeuLysLysProSerLysTyrCysAsnTrpLysCysAlaAlaLeuSer 345
Db 1893 CTTGCCTTCACTTTTAAACAAACCATACAGATGCTACAATTTGGAAGTGCACAGCCTTGAGC 1952
QY 346 AlaIleValIleSerAlaThrLeuValIleLeuLeuAlaTyrPheVal-AlaMethHisLe 365
Db 1953 GCCACTGCAATCACAGTACTTGGCCTTGTACTAGCCTATGTAATTGGTAAGCCCTTT 2012
QY 365 uPhe-----GlyLeuAsnTr 370
Db 2013 ATTCTATGTGTGAGCTTTCTTTTAACTGCCTGCTTTTGTTCAGTGGGTAGGAAT 2072
QY 370 pHisLeuGlnProMetGluGlyGlnMetTyrGluIle 382
Db 2073 TCACCTCAATGTTGTGATGAGCAAAATTTGTAGCATA 2109
RESULT 12
AL922332/c 490 bp mRNA linear EST 18-SEP-2002
LOCUS AL922332 PJR-Z1+Z2 Danio rerio cDNA clone 109-E12-2, mRNA sequence.
DEFINITION AL922332
ACCESSION AL922332
VERSION AL922332.1 GI:23188912
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 490)
AUTHORS Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
Wang,W., Wen,Z. and Peng,J.
15000 unique zebrafish EST clusters and their future use in
microarray for profiling gene expression patterns during
embryogenesis
Genome Res. 13 (3), 455-456 (2003)
MEDLINE 22505427
PUBMED 12618376
COMMENT Contact: Peng J
Lab of Functional Genomics
Institute of Molecular and Cell Biology
30 Medical Drive, Singapore, 117609, Singapore
Email: pengjr@imcb.a-star.edu.sg
Clone requests: pengjr@imcb.a-star.edu.sg.
Location/Qualifiers
FEATURES
source
1. .490
/organism="Danio rerio"
/mol_type="mRNA"
/strain="local wildtype"
/db_xref="taxon:7955"
/clone="109-E12-2"
/tissue_type="whole embryo or fish"
/dev_stage="mixed stages"
/clone_lib="PJR-Z1+Z2"
ORIGIN
Alignment Scores: 8.72e-34 Length: 490
Pred. No.: 596.50 Matches: 116
Score: 82.10% Conservative: 17
Percent Similarity: 71.60% Mismatches: 28
Best Local Similarity: 27.83% Indels: 1
Query Match: 9 Gaps: 1
DB:


```
US-10-029-020-14_COPY_1_400 (1-400) x AL922332 (1-490)
QY      3 VallysGluArgLysProTyrArgSerLeuThrArgArgArgAlaGluArgArgTyr 22
Db      489 GTGAAGGAACGACAGACCTACCGCTCTCTGACCTCCAGGCGGGACACGAGCGCGCTAC 430
QY      23 ThrSerSerAlaAspSerGluGluGlyLysAlaProGlnLysSerTyrSerSerSer 42
Db      429 ACCAGCTCATCTCCGACAGCGAGGATGGCAAGATCAACCCCTAAATCTTACAGTCGAGT 370
QY      43 GluThrLeuLysAlaTyrAspGlnAspAlaArgLeuAlaTyrGlySerArgValLysAsp 62
Db      369 GAGACCTCAAGCCTTCGACCGACTCCAGACTGGCTTATGGCAGCGCGTCAAAGAC 310
QY      63 IleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsnPhetheThrLeuArgGlu 82
Db      309 CTGGTGACCATGAGGCCGACGAGTTCAGCAGACAAGGCCGGACTTTTCTCTCAGAGAC 250
QY      83 LeuGlyLeuGluGluValThrProProHisGlyThrLeuTyrArgThrAspIleGlyLeu 102
Db      249 ATGGCCTTCGGAGATCCCGTCCGCCACATATGGCAGCGTACCGGACAGAAATGGGCCTT 190
QY      103 ProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMetGluAlaAspThrVal 122
Db      189 CCCACCGCGACTACTCAGTGAGTGTGGCATCAGACGCTGACACAGAAACAGACGGCATA 130
QY      123 LeuSerProGluHisProValArgLeuTyrGlyArgSer---ThrArgSerGlyArgSer 141
Db      129 ATGTCCCCCGAGCGCGCTCCGCCCTCTGGGCGCGCAGCAACACCAATCCGCGCGCAGT 70
QY      142 SerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThrGluHisGluAsn 161
Db      69 TCCTGCCTGTCCAGCAGAGCCAACTCCAACTTACGCTCACCAGACACCGAGCATGAAAC 10
QY      162 ThrGlu 163
Db      9 ACTGAG 4

RESULT 13
LOCUS      BG036207
DEFINITION 602326960F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428351 5',
            mRNA sequence.
VERSION     BG036207
KEYWORDS    BG036207.1 GI:124311132
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 870)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10178 row: h column: 16
            High quality sequence stop: 713.
            Location/Qualifiers
              1..870
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4428351"
                /tissue_type="adenocarcinoma, cell line"

FEATURES
source
```

```
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_91"
/note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.4 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
```

ORIGIN

```
Alignment Scores:
Pred. No.:      2.39e-31      Length:      870
Score:          568.00      Matches:    109
Percent Similarity: 99.09%      Conservative: 0
Best Local Similarity: 99.09%      Mismatches: 1
Query Match:      26.50%      Indels:    1
DB:              10          Gaps:      0
```

US-10-029-020-14_COPY_1_400 (1-400) x BG036207 (1-870)

```
QY      291 GlyThrSerProLeuPheCysThrThrSerProGlyTyrProLeuThrSerSerThrVal 310
Db      2 GGCACCTCCCCGCTCTTCTGCACCATCACCAGG-TACCCACTGACGTCCAGCACAGTG 60
QY      311 TyrSerProProArgProLeuProArgSerThrPheAlaArgProAlaPheAsnLeu 330
Db      61 TACTCTCTCCGCCCGACCCCTGCCCGCAGCACCTTGCCTCCGCGCTTAACTC 120
QY      331 LysLysProSerLysTyrCysAsnTyrPheCysAlaAlaLeuSerAlaIleValIleSer 350
Db      121 AAGNAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCCATCGTCATCTCA 180
QY      351 AlaThrLeuValIleLeuLeuAlaTyrPheValAlaMetHisLeuPheGlyLeuAsnTyr 370
Db      181 GCCACTCTGCTCATCTGCTGGCATACTTGTGGCCATGCACCTGTGGCCTAAACTGG 240
QY      371 HisLeuGlnProMetGluGlyGlnMetTyrGluIleThrGluAspThrAlaSerSerTyr 390
Db      241 CACCTGCAGCGGATGGAGGGGCGAGATGTATGATGATCAGGAGGACACAGCCAGTGG 300
QY      391 ProValProThrAspValSerLeuTyrPro 400
Db      301 CCTGTGCCAACCGACGCTCTCCCTATACCC 330
```

RESULT 14

AY405291

LOCUS

DEFINITION Homo sapiens HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

AY405291

KEYWORDS AY405291.1 GI:39761265

SOURCE GSS.

ORGANISM Homo sapiens (human)

Homo sapiens

REFERENCE 1 (bases 1 to 511)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

TITLE

JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 511)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT

This sequence was made by sequencing genomic exons and ordering

Inferring nonneutral evolution from human-chimp-mouse orthologous

TITLE

gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 511)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..511
/organism="pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>511
/locus_tag="HCM2175"

ORIGIN

Alignment Scores: 5.13e-29 Length: 511
Pred. No.: 531.50 Matches: 108
Score: 79.88% Conservative: 23
Percent Similarity: 65.85% Mismatches: 30
Best Local Similarity: 24.80% Indels: 3
Query Match: 29 Gaps: 3
DB:

US-10-029-020-14_COPY_1_400 (1-400) x AY405291 (1-511)

1 MetAspValLysGluArgLysProTyrArgSerLeuThrArg---ArgArgAspAlaGlu 19
1 ATGGATGTGAAGAACGCGAGGCTTACTGCTCCCTGACCAAGACGACGAGAGAGGAA 60
20 ArgArgTyrThrSerSerAlaAspSerGluGluGlyLysAlaPro---GlnLysSer 38
1 CGGCGCTACACAAATTCCTCCGACGACAAATGAGGAGTGCCGGGTACCCACACAGAAGTCC 120
39 TyrSerSerSerGluThrLeuLysAlaTyrAspGlnAsp---AlaArgLeuAlaTyrGly 57
121 TACAGTTCACGAGACATTTGAAAGCTTTTGATCATGATTCCTCGCGGCTGCTTTACGGC 180
58 SerArgValLysAspIleValProGlnGluAlaGluGluPheCysArgThrGlyAlaAsn 77
181 AACAGAGTGAAGATTTGGTTCCACAGAGAAGCAGAGAGTTCACTAGACAAGGACAGAAT 240
78 PheThrLeuArgGluLeuGlyLeuGluGluValThrProHisGlyThrLeuTyrArg 97
241 TTTACCTTAAGGAGTAGGAGTTGTGTAACCAACCACTCGAAGAGGACTGGCATTTTGT 300
98 ThrAspIleGlyLeuProGlnCysGlyTyrSerMetGlyAlaGlySerAspAlaAspMet 117
301 GCGGAATGAGCAGTGTATGTCCTCCAGAGCATGCCATGAGACTTTGGGGCAGGGGTCAAA 420
118 GluAlaAspThrValLeuSerProGluHisProValArgLeuTrpGlyArgSerThrArg 137
361 GAAATGAAGCAGTGTATGTCCTCCAGAGCATGCCATGAGACTTTGGGGCAGGGGTCAAA 420
138 SerGlyArgSerSerCysLeuSerSerArgAlaAsnSerAsnLeuThrLeuThrAspThr 157
421 TCAGGCGCAGCTCCTGCCTGTCAAGTCGGTCCCACTCAGCTCAGCCCTCACCATGACGATACG 480
158 GluHisGluAsn 161
481 GAGCACGAAAAAC 492

RESULT 15
AY405292
LOCUS
DEFINITION
Pan troglodytes HCM2175 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY405292
VERSION
AY405292.1 GI:39761266
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 511)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.

Search completed: August 14, 2004, 18:02:37
Job time : 4710.12 secs